AstraZeneca

Statistical Analysis Plan

Study Code	D081DC00007
Edition Number	4.0
Date	4 July 2019

A Phase III, Open Label, Randomized Study to Assess the Efficacy and Safety of Olaparib (Lynparza™) Versus Enzalutamide or Abiraterone Acetate in Men with Metastatic Castration-Resistant Prostate Cancer Who Have Failed Prior Treatment with a New Hormonal Agent and Have Homologous Recombination Repair Gene Mutations (PROfound)

A Phase III, Open Label, Randomized Study to Assess the Efficacy and Safety of Olaparib (Lynparza™) Versus Enzalutamide or Abiraterone Acetate in Men with Metastatic Castration-Resistant Prostate Cancer Who Have Failed Prior Treatment with a New Hormonal Agent and Have Homologous Recombination Repair Gene Mutations (PROfound)

[redacted]

[redacted]

Date

A Phase III, Open Label, Randomized Study to Assess the Efficacy and Safety of Olaparib (Lynparza™) Versus Enzalutamide or Abiraterone Acetate in Men with Metastatic Castration-Resistant Prostate Cancer Who Have Failed Prior Treatment with a New Hormonal Agent and Have Homologous Recombination Repair Gene Mutations (PROfound)

[redacted]

[redacted]

Date

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
ALT	Alanine aminotransferase
AQA	Analgesic Quantification Algorithm
ATM	Ataxia telangiectasia mutated
AST	Aspartate aminotransferase
BICR	Blinded Independent Central Review
bid	Twice daily (Latin: *bis die*)
BOR	Best objective response
BP	Blood pressure
BPI-SF	Brief Pain Inventory – Short Form
BRCA	Breast cancer gene, i.e., BRCA1 and BRCA2
CI	Confidence interval
CR	Complete response
CRPC	Castration-resistant prostate cancer
CSP	Clinical study protocol
CSR	Clinical Study Report
CT	Computed tomography
CTC	Circulating tumor cell
CTCAE	Common Terminology Criteria for Adverse Event
CV	Coefficient of variation
DAE	Discontinuation of Investigational Product due to Adverse Event
DCO	Data cut-off
DNA	Deoxyribonucleic acid
DOR	Duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EFR	Evaluable for response
EQ-5D-5L	EuroQuol-5 Dimensions, five level
EWB	Emotional well-being

Abbreviation or special term	Explanation
FACIT	Functional assessment of chronic illness
FACT-G	Functional Assessment of Cancer Therapy - General
FACT-P	Functional Assessment of Cancer Therapy - Prostate Cancer
FAPSI-6	FACT Advanced Prostate Symptom Index 6
FAS	Full analysis set
FMI	Foundation Medicine Inc.
FWB	Functional well-being
HRR	Homologous recombination repair
HR	Hazard ratio
HRQL	Health-Related Quality of Life
ICU	Intensive care unit
IDMC	Independent Data Monitoring Committee
IP	Investigational product
IPCW	Inverse Probability of Censoring Weighting
IPD	Important protocol deviation
KM	Kaplan-Meier
LD	Longest diameter
LLOQ	Lower limit of quantification
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed model for repeated measures
MRI	Magnetic resonance imaging
MTP	Multiple testing procedure
NA	Not applicable
NC	Not calculable
NE	Not-evaluable
NED	No evidence of disease
NHA	New hormonal agent
NQ	Non-quantifiable
NR	Not Reportable
NS	No Sample
NTL	Non-target lesion
OAE	Other Significant Adverse Event

Abbreviation or special term	Explanation
OME	Oral morphine equivalence
ORR	Objective response rate
OS	Overall survival
PARP	Polyadenosine 5'-diphosphoribose polymerase
PCS	Prostate cancer subscale
PCWG-3	Prostate Cancer Working Group 3
PD	Progressive disease
PFS2	Second progression
PGIC	Patient Global Impression of Change
PID	Percentage intended dose
PK	Pharmacokinetics
PR	Partial response
PRO	Patient Reported Outcome
PSA	Prostate specific antigen
PWB	Physical well-being
RDI	Relative dose intensity
RECIST 1.1	Response Evaluation Criteria in Solid Tumors version 1.1
rPFS	Radiologic progression-free survival
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Stable disease
SOC	System organ class
SSRE	Symptomatic Skeletal –Related Event
SWB	Social well-being
TL	Target lesion
TOI	Trial outcome index
VAS	Visual analogue scale

AMENDMENT HISTORY

Date	Brief description of change
4 July 2019	SAP v4.0 vs SAP v3.0 • Clarified concomitant medications will be summarized for Cohort A FAS and Cohort B FAS in addition to Cohort A+B FAS • Updated rPFS censoring to censor at the time of the earliest date of their last evaluable RECIST 1.1 assessment (taking the latest target lesion, non-target lesion or new lesion scan date) or bone scan assessment that showed Non-PD. The definition now uses the bone scan visit response description rather than the bone scan visit definition. • Section 3.2.2 Analgesic Use Scoring - Moved AQA missing data imputation rules from TTPP section to this section - Added additional information for the imputation rules for OME and AQA scores. If "other" medications are clearly identified as non-opioid the OME score will be set to 0. - Added additional information for AQA score imputation. If additional pain medications taken alongside "Other" over the 7 days of assessments, and they are all non-opioids, then AQA score of 1 will be assigned (i.e. non-opioid analgesics) • Moved Time to pain progression section to Section 3.3.3 (previously Section 3.3.4) • Section 3.3.3 Time to pain progression (TTPP): - Removed the text which required the 2 consecutive subsequent assessments to be separated by 3-4 weeks. The requirement is 2 consecutive follow-up assessments (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit). - Added additional information regarding how the average BPI-SF worst pain [Item 3] and the average AQA score are derived - Added additional information regarding baseline, assessments on or before the date of first treatment will be considered screening - Added definition of a missed visit for pain progression - Added additional clarity to show that olaparib is a considered a subsequent therapy for patients who switch from investigators choice of NHA to olaparib upon progression - Added additional information for the censoring rules for time to pain progression • Section 3.3.4 Brief Pain Inventory – short form (BPI-SF): - Updated pain severity subscale/domain to include average pain severity definition - Removed the text which required the 2 consecutive subsequent assessments to be separated by 3-4 weeks. The requirement is 2

Date	Brief description of change
	consecutive follow-up assessments (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit). - Clarified that the date of pain severity progression will be the earliest of the assessments contributing to the average. - Added additional information regarding how the average BPI-SF "pain severity" subscale score and the average AQA score are derived - Added additional information regarding baseline, assessments on or before the date of first treatment will be considered screening - Added definition of a missed visit for pain severity progression - Added additional clarity to show that olaparib is a considered a subsequent therapy for patients who switch from investigators choice of NHA to olaparib upon progression - Added additional information for the censoring rules for time to pain severity progression • Section 3.3.5 Pain Palliation: - Updated pain severity subscale/domain to include average pain severity definition - Clarified pain palliation uses the average BPI-SF worst pain [Item 3] - Updated confirmation visit definition to clarify start date of subsequent visit – start date of initial visit must be >= 14 days - Added additional information regarding how the average BPI-SF worst pain [Item 3] and the average AQA score are derived - Added additional information regarding baseline, assessments on or before the date of first treatment will be considered screening - Added definition of a missed visit for pain palliation - Added additional clarity to show that olaparib is a considered a subsequent therapy for patients who switch from investigators choice of NHA to olaparib upon progression • Updated FACT-P TOI worsening examples to reflect the criteria in Table 12 • Clarified the p-values reported for the interaction testing subgroups will represent those from the final model resulting from stepwise backwards selection; the 'selection model' • Removed the cumulative incidence function from the cumulative incidence plot for AESIs in Section 4.2.5.1 Adverse events • Updated Section 6 Changes of Analysis from Protocol
29 April 2019	SAP v3.0 vs SAP 2.0 • [redacted] • Added clarification to the origin (eCRF or IVRS) of measurable disease

Date	Brief description of change
	and prior taxane use throughout • Changed 'Investigator choice' to 'Investigator choice of NHA' throughout • Added clarity throughout for rPFS, sensitivity analysis and all secondary endpoints to show they will all use the same strata as the primary rPFS analysis as determined by the pooling strategy • Specified that there needs to be at least 5 responses to perform logistic regression analyses throughout, otherwise a fisher's exact test will be used • Added text throughout for PRO endpoints to show analyses will include data until the start date of subsequent anti-cancer therapy for patients who receive a subsequent anti-cancer therapy (note that radiotherapy is not considered a subsequent anti-cancer therapy). • Section 2.1 (Definition of analysis sets) - Clarified ORR, DoR and BoR will use the EFR analysis set - Added EFR analysis set for Cohort A, B and A+B - Update to Table 6 to reflect the addition of EFR analysis set • Added IPD for patients who were randomized to investigators choice of NHA and started olaparib treatment before disease progression determined by BICR • Removed partial date imputation rule from Section 3 and added to Section 3.3.6 (Overall survival) only • Update to Table 9 – overall visit soft tissue responses • Removed BICR adjudication details from SAP, this can be found in imaging charter • Clarified BICR will be carried out on PCWG-3 assessments too • Section 3.2 (rPFS) - Update to rPFS censoring approach for censoring patients who have not progressed or died at the time of analysis and for censoring patients who progress or die immediately after 2 or more consecutive missed visits. The updated approach will take the earliest date of their last evaluable RECIST 1.1 assessment (taking the latest target lesion, non-target lesion or new lesion scan date) or bone scan assessment that showed fewer than 2 new lesions - Update to Table 11 – overall radiological visit response - Text added describing how RECIST 1.1 and PCWG3 assessments will be merged for BICR and investigator assessments • Updated ORR to use the EFR analysis set • Section 3.3.3 (BPI-SF) - Added clarification to regarding visits separated by 3-4 weeks - Added definition of completed questionnaire to PRO compliance - Text added regarding completion rate definition and PRO patient

Date	Brief description of change
	disposition tables. • Section 3.3.4 (Time to pain progression) - Added clarification regarding visits separated by 3-4 weeks - Added text to show the date of pain progression is the earliest date when the average is taken (i.e. average of 7-day assessments) "worst pain in 24 hours" (BPI-SF item 3) - Added missing data imputation rule for AQA • Added clarification regarding visits separated by 3-4 weeks to section 3.3.5 (Pain palliation) • Clarified survival calls will be made for the primary rPFS analysis, final OS analysis [redacted] • Section 3.3.8 (Duration of response) - Update to DoR to use the EFR analysis set - Update to censoring rule - Added clarification to time to response definition - Unconfirmed response definition added in • Clarified which analysis set will be used for PSA response, PSA changes on a continuous scale, CTC conversion rate and CTC counts on a continuous scale • Added text to show patients data will only be included until the start date of the subsequent anti-cancer therapy for patients who receive a subsequent anti-cancer therapy for PSA response, PSA changes on a continuous scale, CTC conversion rate and CTC counts on a continuous scale • Section 3.3.14 (FACT-P) - Clarified that improvement rates and time to deterioration will be performed for FACT-G - Clarification to Table 13 to show consecutive visits need to be at least 3 weeks apart - Clarification to time to deterioration to show there must be no improvement between subsequent visits - Removed reference to 'patient is too affected by symptoms' • Text added to general considerations for safety regarding visit windows and methods for handing data in visit based summaries • Removed compliance definition and summary • Clarified exposure will not be calculated for prednisone/prednisolone • Added percentage intended dose definition and summary • Update to Table 14 to include all planned analyses and cohorts • Section 4.2.2 - Clarified that the pooling strategy will be employed for Cohort A, Cohort B and Cohort A+B separately

Date	Brief description of change
	- Added text to show all sensitivity analyses and secondary endpoints (except for ORR which only includes prior taxane) will use the same strata as the primary model - Clarified that the rPFS analysis and KM plot will be produced for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients • Section 4.2.2.1 (Subgroup analysis) - Update to subgroup analysis to only provide descriptive statistics if there are less than 5 events across both treatment groups - Clarification added to stratification factors - Clarification added to show which HRR gene mutation subgroups will be performed in the FAS, confirmed F1CDx patients and confirmed myriad *gBRCA*m patients - Clarified that the results for HRR gene mutation subgroups will be displayed for one level only - Removed germline vs somatic subgroup - Added KM plots for selected subgroups - Added text to show that the consistency of treatment effect between subgroups will not use HRR gene mutation subgroups • Section 4.2.2.2 (Sensitivity analysis) - Added information for previous evaluable assessment to select sensitivity analyses - Added KM plot to censoring bias and ascertainment bias sensitivity analysis - Clarified how unequivocal clinical progression is determined • Removed text which stated that for the key secondary analyses, the subgroup analyses specified for the primary analysis and the sensitivity analyses for HRR mutation status will be repeated • Section 4.2.3.1 (Confirmed objective response rate) - Added BOR analysis on soft tissue using RECIST 1.1 - Added all ORR analysis will be repeated for patients with an unconfirmed response - Added text to show that the confirmed ORR logistic regression will be produced for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients - Added subgroup analysis for confirmed ORR • Added logistic regression for pain progression, pain severity, pain palliation • Added text to show analysis of pain progression, pain severity and pain palliation will be repeated for patients who are non-opiate users at baseline • Added text to show TTPP will be produced for confirmed FMI F1CDx

Date	Brief description of change
	patients and confirmed myriad *gBRCA*m patients • Added missing data summary and listing for time to pain progression • Clarified that the OS analysis, not including the KM plot, will be produced for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients • Added DoR analysis for patients with an unconfirmed response • Clarified that PSA waterfall plots will be created for best percentage change and percentage change at Week 12 • Added visit window around the Week 12 visit for PSA • Clarified CTC waterfall plots and summaries will be produced for best change from baseline and best percentage change from baseline • Updated text to show FACT-P scores will be presented by treatment group for all visits until there are less than the minimum of 20 or 1/3 of patients dosed • Update to FACT-P to use the logistic regression methods described for TTPP • Added prior taxane and measurable disease at baseline into the MMRM model for FACT-P • Removed text 'The population PK/pharmacodynamic modelling will be reported separately from the CSR.' • Clarified that the overall summary AE table will be produced for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients • Updated concomitant and disallowed medications to be summarized using the FAS • Added summary for disallowed medications • Removed concomitant medication summary for patients who switch from investigators choice of NHA to olaparib • Added text to show PRO-CTCAE will be summarized for only patients in countries (Argentina Australia, Austria, Canada [English-speaking sites], Germany, Japan, Spain, United Kingdom and United States) where the questionnaire was administered • Clarified which demographic and baseline data will be summarised for the FAS, confirmed FMI F1CDx patients, confirmed myriad gBRCAm and the patients who switch from investigators choice of NHA to olaparib
09 November 2018	SAP v2.0 vs SAP v1.0 • Replaced the use of 'subjects' with 'patients' throughout • Added additional Exploratory objectives for ctDNA to reflect the clinical study protocol (CSP) • Updated Full Analysis Set definitions to explicitly describe by cohort

Date	Brief description of change
	• Added Safety switch analysis set • Updated Table 6 to contain all primary and secondary endpoints with analysis populations • Section 2.2 (Violations and deviations): - Removed repeated important protocol deviations - Text relating to mis-randomizations in treatment dispensing removed since these are covered within retained list of important protocol deviations - Added text describing a deviation bias sensitivity analysis • Added partial date imputation rules to primary and secondary variables section • New sections added to section 3.1, including Target lesions, Non-target lesions, Overall visit response and Bone lesion progression using PCWG3 • Section 3.2 (Primary endpoint- Radiological Progression Free Survival [rPFS]): - Updated definition of radiological progression-free survival to also refer to censoring - Updated definition of censored patients to include patients who have not died at time of analysis - Added detail for allowable intervals for 8 weekly scheduled scans • Section 3.3.1 (Confirmed Overall Objective Response Rate [ORR]) - Added definition of an unconfirmed response - Added rules for non-consecutive visits • Added section 3.2.2 for Analgesic Use Scoring • Section 3.3.3 (Brief Pain Inventory – short form [BPI-SF]): - Clarified all items must be non-missing for pain severity score - Time to pain progression and pain palliation moved to sections 3.3.4 and 3.3.5 • Calculation for overall survival calculation added in for clarification • Section 3.3.8 (Duration of Response [DOR]): - Added text to DOR definition clarifying patients must have measurable disease at baseline and have a confirmed response - Added DOR calculation - Added text to show response must be confirmed • Added section 3.3.10 (Time to Opiate Use for Cancer Pain) • Text added regarding missing PSA data • Added section 3.3.11 (PSA changes on continuous scale) • Text added to clarify CTC conversion rate can be at any visit post baseline

Date	Brief description of change
	• Text added regarding missing CTC values • Detail added concerning second progression • Section 3.3.14 (Functional Assessment of Cancer Therapy- Prostate Cancer [FACT-P]): - Added FACT-G score throughout - Added clarification for missing data in FACT-G and TOI subscales - Updated the definition of visit response in Table 12 - Updated text for time to deterioration for FACT-P • Added text for length of hospital stay and length of ICU stay • Added Section 3.5.1 (General considerations for safety assessments) and Section 3.5.2 (Handling of partial dates) • Clarified definition of treatment emergent AEs • Updated concomitant medication coding dictionary to WHO drug dictionary • Added exposure calculation for patients who switch • Added detail to Compliance and Exposure regarding Missed or Forgotten doses and Safety Follow-up • Updated baseline definition for PRO endpoints • Section 4.2.2 (Analysis of the primary efficacy variable [rPFS]): - Updated to show the hypothesis of superiority of olaparib compared to investigator choice will be tested using a log rank test, and removed reference to p-value from Cox Proportional Hazards Model - Updated pooling strategy - Added detail to the assessment of proportionality • Section 4.2.2.1 (Subgroup analysis): - Added Cohort A+B - Detailed the HRR gene mutation combinations - Added additional subgroups • Removed sensitivity analysis for enrichment (h) since no enrichment was implemented • Added text to repeat HRR mutation subgroup analysis for key secondary endpoints • Added text for analysis of soft tissue ORR and text for all ORR analyses to be repeated using investigator assessed response • Clarified the subgroup analysis detailed for rPFS will be repeated for overall survival at the final analysis only • Added Section 4.2.3.5 (Duration of Response) and Section 4.2.3.6 (Time to Opiate use for Cancer related Pain) • Added text for percentage change from baseline to be summarized for

Date	Brief description of change
	CTC counts • Added supportive analyses for FACT-P • Added detail to pharmacokinetic analysis section regarding the summary table and listing • Added Section 4.2.4 (Concordance between BICR and investigator assessments) • Corrected criteria for death summaries that display deaths > 30 days after last treatment dose • Removed Section 4.2.4.2 (Dose limiting toxicities) • Text added to Section 4.2.5.1 (Adverse events) regarding a new listing and separate summaries for patients who switch • Added shift table to Section 4.2.5.2 (Laboratory assessments) for patients who switch • Section 4.2.5.3 (Concomitant medications): - Updated text to clarify summaries will be for Cohort A+B only - Added summary for patients who switch from investigators choice of treatment to olaparib • Added vital signs summary for patients who switch from investigators choice of treatment to olaparib • Section 4.2.5.5 (Compliance and exposure): - Updated text to clarify summaries will be for Cohort A+B only - Added text relating to summaries for subset of patients with any qualifying HRR mutation by the FMI central tumor test - Added text relating to summaries for patients who switch from investigators choice of treatment to olaparib • Section 4.2.6 (Supportive analysis): - Updated text to clarify that resource use and EQ 5D 5L may be reported outside of the CSR - Added AQA and FACT-G to PRO summaries - Added absolute and change from baseline mean (± standard deviation) plots for FACT-P total score, FACT-G total score, TOI, FAPSI-6, PCS, FWB, PWB, EWB, SWB, BPI-SF item #3 (worst pain in 24 hours), pain severity (BPI-SF pain severity domain), pain interference and AQA score - Added additional supportive analyses for BPI-SF (BPI-SF Item 3, pain severity score and pain interference score) - Added text detailing which cohorts will be presented for PRO endpoints and that the analysis will not be performed for patients who switch from investigators choice to olaparib • Updated list of demographic and baseline data that will be listed to meet ICH requirements • Added summaries to Section 4.2.7 (Demographic and baseline data) for

Date	Brief description of change
	patients who switch from investigator choice of treatment to olaparib • Added text to Section 6 (Changes of Analysis from Protocol) based on the above

1. STUDY DETAILS

1.1 Study objectives

1.1.1 Primary Objective

Table 1 Primary objectives

Primary objective	Primary outcome measures
To determine the efficacy (as assessed by radiographic Progression- Free Survival (rPFS)) of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with metastatic Castration-Resistant Prostate Cancer (mCRPC) with *BRCA1*, *BRCA2* or *ATM* qualifying mutations (Cohort A)	• rPFS by blinded independent central review（BICR) using Response Evaluation Criteria in Solid Tumors (RECIST 1.1, soft tissue) and Prostate Cancer Working Group 3 (PCWG3, bone) criteria

1.1.2 Secondary Objectives

Table 2 Key secondary objectives

Key Secondary Objectives	Key secondary outcome measures
To determine the efficacy (as assessed by Objective Response Rate［ORR]) of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with *BRCA1*, *BRCA2* or *ATM* qualifying gene mutations (Cohort A) [redacted]	• Confirmed ORR by BICR assessment in patients with measurable disease at baseline using RECIST 1.1 (soft tissue) and PCWG3 (bone) criteria
To determine the efficacy (as assessed by rPFS) of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with Homologous Recombination Repair (HRR) qualifying mutations (Cohort A+B)	• rPFS by BICR using RECIST 1.1 (soft tissue) and PCWG3 (bone) criteria
To determine the efficacy (as assessed by time to pain progression) of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with *BRCA1*, *BRCA2* or *ATM* qualifying gene mutations (Cohort A)	• Pain progression based on Brief Pain Inventory- Short Form (BPI-SF) item 3 “worst pain in 24 hours” and opiate analgesic use (AQA score)

Table 2 Key secondary objectives

Key Secondary Objectives	Key secondary outcome measures
To determine the efficacy (as assessed by overall survival) of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with *BRCA1*, *BRCA2* or *ATM* qualifying gene mutations (Cohort A)	• Overall survival (OS)

Table 3 Other secondary objectives

Other Secondary Objectives	Other secondary outcome measures
To further assess the efficacy of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with *BRCA1*, *BRCA2* or *ATM* qualifying gene mutations (Cohort A)	• Time from randomization to the first Symptomatic Skeletal – Related Event (SSRE) • Time from partial or complete response in patients with measurable disease at baseline (RECIST 1.1) to progression by BICR (Duration of Response [DoR]) • Time from randomization to opiate use for cancer-related pain • Confirmed ORR (RECIST 1.1) in soft tissue by BICR in patients with measurable disease at baseline (Soft tissue response) • Proportion of Patients achieving a ≥50% decrease in Prostate Specific Antigen (PSA) from baseline to the lowest post-baseline PSA result, confirmed by a second consecutive PSA assessment at least 3 weeks later (PSA_{50} response) • Proportion of Patients achieving a decline in the number of Circulating Tumor Cells (CTCs) from ≥ 5 cells/7.5mL to < 5 cells/7.5mL whole blood (CTC conversion rate) • Time from randomization to second progression by investigator assessment of radiological or clinical progression or death (PFS2)

Table 3 Other secondary objectives

Other Secondary Objectives	Other secondary outcome measures
To further assess the effect of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with *BRCA1*, *BRCA2* or *ATM* qualifying gene mutations (Cohort A) on disease-related symptoms and Health-Related Quality of Life (HRQoL)	• Pain severity progression based on BPI-SF Pain Severity domain/subscale and opiate analgesic use (AQA score) • Pain interference based on BPI-SF Pain Interference domain/subscale • Functional Assessment of Cancer Therapy- Prostate cancer [FACT-P] (FACT-P total score, Trial Outcome Index [TOI], Functional Well-Being [FWB], Physical Well-Being [PWB], Prostate Cancer Subscale [PCS], and FACT Advanced Prostate Symptom Index 6 [FAPSI 6]) • Proportion of Patients with pain (BPI-SF item 3) score ≥ 4 points at baseline who have a decrease of ≥ 2 points in pain (BPI-SF item 3) and without ≥ 1 point increase in analgesic score (AQA score) at 12 weeks, confirmed at least 3 weeks later (Pain palliation)
To assess the efficacy of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with HRR qualifying gene mutations other than *BRCA1*, *BRCA2* or *ATM* (Cohort B) [redacted]	• rPFS by BICR using RECIST 1.1 (soft tissue) and PCWG3 (bone) criteria • Confirmed ORR by BICR assessment in patients with measurable disease at baseline using RECIST 1.1 (soft tissue) and PCWG3 (bone) criteria • Pain progression based on BPI-SF item 3 "worst pain in 24 hours" and opiate analgesic use (AQA score) • OS

Table 3 Other secondary objectives

Other Secondary Objectives	Other secondary outcome measures
To further assess the efficacy of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with HRR qualifying gene mutations (Cohort A+B) [redacted]	• Confirmed ORR by BICR assessment in patients with measurable disease at baseline using RECIST 1.1 (soft tissue) and PCWG3 (bone) criteria • Time from randomization to the first SSRE • Time from partial or complete response in patients with measurable disease at baseline (RECIST 1.1) to progression by BICR (DoR) • Time from randomization to opiate use for cancer-related pain • Confirmed ORR (RECIST 1.1) in soft tissue by BICR in patients with measurable disease at baseline (Soft tissue response) • Proportion of Patients achieving a ≥50% decrease in PSA from baseline to the lowest post-baseline PSA result, confirmed by a second consecutive PSA assessment at least 3 weeks later (PSA_{50} response) • Proportion of Patients achieving a decline in the number of CTCs from ≥ 5 cells/7.5mL to < 5 cells/7.5mL whole blood (CTC conversion rate) • Time from randomization to second progression by investigator assessment of radiological or clinical progression or death (PFS2) • OS

Table 3 Other secondary objectives

Other Secondary Objectives	Other secondary outcome measures
To further assess the effect of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with HRR qualifying gene mutations (Cohort A+B) on disease-related symptoms and Health-Related Quality of Life (HRQoL) [redacted]	• Pain progression based on BPI-SF item 3 "worst pain in 24 hours" and opiate analgesic use (AQA score) • Pain severity progression based on BPI-SF Pain Severity domain/subscale and opiate analgesic use (AQA score) • Pain interference based on BPI-SF Pain Interference domain/subscale • FACT-P (FACT-P total score, Trial Outcome Index [TOI], Functional Well-Being [FWB], Physical Well-Being [PWB], Prostate Cancer Subscale [PCS] and FACT Advanced Prostate Symptom Index 6 [FAPSI 6]) • Proportion of Patients with pain (BPI-SF item 3) score ≥ 4 points at baseline who have a decrease of ≥ 2 points in pain (BPI-SF item 3) and without ≥ 1 point increase in analgesic score (AQA score) at 12 weeks, confirmed at least 3 weeks later (Pain palliation)
To determine the exposure to olaparib in a subset of patients receiving olaparib [redacted]	• Olaparib plasma concentration data

1.1.3 Safety Objectives

Table 4 Safety objectives

Safety objective	Safety outcome measures
To evaluate the safety and tolerability of olaparib versus investigator choice of enzalutamide or abiraterone acetate (Cohort A+B)	• Adverse events (AEs)/ serious adverse events (SAEs) • Collection of clinical chemistry/haematology parameters

1.1.4 Exploratory Objectives

These exploratory objective analyses may be reported separately from the clinical study report.

Table 5 Exploratory objectives

Exploratory objective	Exploratory outcome measures
To compare the effect of olaparib versus investigator choice of enzalutamide or abiraterone acetate on patient-reported treatment tolerability and overall health status (Cohort A)	• Patient Reported Outcomes-Common Terminology Criteria for Adverse Events (PRO-CTCAE) • Patient Global Impression of Change (PGIC)
To compare the effect of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with *BRCA1*, *BRCA2* or *ATM* qualifying mutations (Cohort A) based on prior receipt of taxane	• Subgroup analysis of rPFS in patients with or without prior taxanes
To compare the effect of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with either germline or somatic *BRCA1*, *BRCA2* or *ATM* qualifying mutations (Cohort A)	• Subgroup analysis of rPFS by BICR based on whether the qualifying mutation is a germline mutation or only in the tumor (somatic)
To compare the effect of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with *BRCA1*, *BRCA2*, or *ATM* qualifying mutations as detected by ctDNA analysis [redacted]	• rPFS analysis in patients with qualifying mutation identified by ctDNA test

Table 5 Exploratory objectives

Exploratory objective	Exploratory outcome measures
To compare the effect of olaparib versus investigator choice of enzalutamide or abiraterone acetate in patients with HRR qualifying mutations as detected by ctDNA analysis [redacted]	• rPFS analysis in patients with qualifying mutation identified by ctDNA test
To explore methods of estimating OS adjusting for the impact of the control arm receiving subsequent PARP inhibitors (including olaparib), platinum compounds or imbalances between the treatment arms for other potentially active agents (Cohort A)	• OS adjusted for impact of subsequent PARP inhibitors (or other potentially active investigational agents)
To compare the tumor HRR gene mutation status in all screened patients with evaluable results from plasma (Cohort A+B) [redacted]	• Comparison of HRR gene mutation status between tumor DNA and plasma derived ctDNA
Future exploratory research into factors that may influence development of cancer and/or response to study treatment (where response is defined broadly to include efficacy, tolerability or safety) may be performed on the collected and stored archival tumor samples that were mandatory for entry onto the study or on blood samples (Cohort A+B) [redacted]	• Evaluate loss of heterozygosity of HRR genes in tumors • Evaluation of ctDNA collected from plasma at baseline and at progression • CTCs (EPIC assay)
To collect and store DNA (according to each country's local and ethical procedures) for future exploratory research into genes/genetic variation that may influence response (i.e., distribution, safety, tolerability and efficacy) to study treatments and or susceptibility to disease (optional) (Cohort A+B)	• Blood sample pharmacogenetics analysis

Table 5 Exploratory objectives

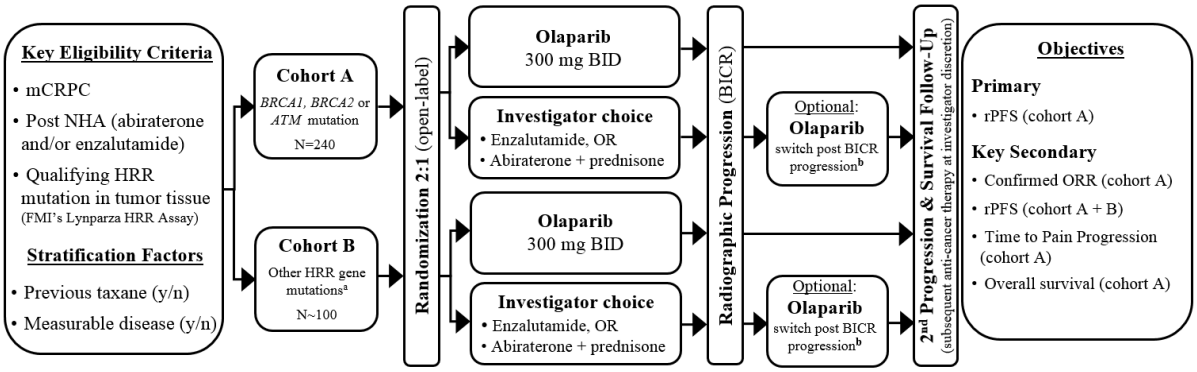
Exploratory objective	Exploratory outcome measures
To investigate the health economic impact of treatment and the disease on hospital related resource use and health state utility (Cohort A)	• Number, type and reason of hospitalizations and hospital attendances, procedures conducted and hospital length of stay (HOSPAD) • EuroQol 5-dimension, 5-level health state utility index (EQ-5D-5L)

Some of these exploratory analyses may be reported separately from the CSR.

1.2 Study design

This is a prospective, multi-centre, open-label, randomized phase 3 trial evaluating the efficacy and safety of olaparib (300 mg orally bid) versus investigator choice of either enzalutamide (160 mg orally od) or abiraterone acetate (1,000 mg orally od with 5 mg bid prednisone) in approximately 340 patients with metastatic castration-resistant prostate cancer (mCRPC) with qualifying homologous recombination repair (HRR) gene mutations who have failed prior treatment with a new hormonal agent (NHA). [redacted] Prior to randomization, it was anticipated that approximately 5500 patients will be centrally screened for eligibility based on tumor tissue testing using the Lynparza HRR assay (Foundation Medicine Inc. (FMI), Cambridge MA). Patients with *BRCA1*, *BRCA2* or *ATM* qualifying gene mutations will be included in Cohort A whereas patients with mutations among 12 other genes involved in HRR (*BARD1, BRIP1, CDK12, CHEK1, CHEK2, FANCL, PALB2, PPP2R2A, RAD51B, RAD51C, RAD51D,* and *RAD54L*) will be included in Cohort B. Patients will be randomized 2:1 to either olaparib or pre-declared investigator choice of either enzalutamide or abiraterone acetate in each of the Cohorts A and B. Randomization will be stratified based on prior receipt of taxane chemotherapy (yes vs no) and presence of measurable disease at baseline (yes vs no). Figure 1 illustrates the overall study design. The global recruitment to the study closed when approximately 340 patients were randomized in Cohorts A and B.

Figure 1 Overall study design



[a] Cohort B HRR genes include *BARD1, BRIP1, CDK12, CHEK1, CHEK2, FANCL, PALB2, PPP2R2A, RAD51B, RAD51C, RAD51D, RAD54L.*
[b] Subjects randomized to investigator choice arm will be given the opportunity to begin treatment with open-label olaparib (300 mg bid) only after objective radiographic progression by blinded independent central reader (BICR). No intervening systemic anti-cancer therapy following discontinuation of randomized treatment will be permitted. Subjects may continue on olaparib as long as they show clinical benefit as judged by the investigator.

[redacted]

1.3 Number of patients

The sample size for Cohort A was based on the research hypothesis that single agent olaparib at 300 mg bid has superior efficacy and an acceptable tolerability profile as compared with enzalutamide or abiraterone acetate in mCRPC patients with deleterious or suspected deleterious HRR gene mutations and who have previously failed treatment with an NHA such as enzalutamide or abiraterone acetate.

Approximately 240 patients will be randomized in Cohort A in a 2:1 ratio to olaparib tablets (300 mg orally bid) versus pre-declared investigator choice of either enzalutamide (160 mg orally od) or abiraterone acetate (1,000 mg orally od with 5 mg bid prednisone). The primary endpoint of the study is rPFS as assessed by BICR.

It is expected that the targeted sample size of 240 patients in Cohort A with approximately 143 rPFS events (60% maturity) will provide 95% power to demonstrate a statistically significant difference in rPFS at a 2-sided alpha level of 5% assuming the true treatment effect was a hazard ratio (HR)=0.53. This translates to an approximately 4.5 month improvement in median rPFS over an assumed 5 month median rPFS on enzalutamide or abiraterone acetate assuming rPFS is exponentially distributed. The smallest treatment difference that would be statistically significant at the final analysis is a HR of 0.71. It is anticipated that the study accrual period will be approximately 28 months and that 143 progression and death events will occur approximately 35 months after the first patient is randomized in the study.

To predict the data cut-off date (DCO) when 143 rPFS events will be first observed in Cohort A, the blinded grouped time to event data will be modelled on an ongoing basis through the trial. Once the DCO is set, data will be collected and cleaned while recognised that there may ultimately be fewer or more events than specified.

Cohort B of the study will consist of approximately 100 patients with qualifying HRR mutations other than *BRCA1*, *BRCA2* and *ATM* as assessed by the FMI central tumor test. These patients will be randomized in a 2:1 ratio to olaparib tablets versus predeclared investigator choice of either enzalutamide or abiraterone acetate.

2. ANALYSIS SETS

2.1 Definition of analysis sets

2.1.1 Full analysis set

Cohort A Full Analysis Set (Cohort A FAS):

The primary statistical analysis of the efficacy of olaparib in comparison to investigator choice of either enzalutamide or abiraterone acetate in Cohort A, will include all patients who were randomized in Cohort A as part of the global enrollment regardless of the treatment actually received. Patients who were randomized in Cohort A as part of the global enrollment but did not subsequently go on to receive study treatment are included in the full analysis set (FAS) on Cohort A. Efficacy and health- related quality of life (HRQoL) data (except for ORR, DoR and BoR) will be analyzed using the full analysis set. See Table 6 for details.

Cohort B Full Analysis Set (Cohort B FAS):

The analysis of the efficacy of olaparib in Cohort B will include all patients randomized to olaparib or investigator choice of either enzalutamide or abiraterone acetate in Cohort B as part of the global enrolment regardless of the treatment actually received. Patients who were randomized in Cohort B but did not subsequently go on to receive study treatment are included in the FAS on Cohort B. Efficacy and HRQoL data (except for ORR, DoR and BoR) will be analyzed using the full analysis set.

Cohort A+B Full Analysis Set (Cohort A+B FAS):

The analysis based on Cohort A+B will include patients from both Cohort A full analysis set and Cohort B full analysis set.

2.1.2 Evaluable for response (EFR) analysis set

This is a subset of the FAS, who have measurable disease at baseline as per the RECIST 1.1 criteria. Measurable disease will be defined using the BICR assessment for analyses of BICR data, as well as using the investigator assessment data for analyses of investigator assessment The EFR set will be defined for Cohort A, Cohort B and Cohort A+B separately. ORR, DoR and BoR will be analyzed using the EFR set.

2.1.3 Safety analysis set

All patients who were randomized as part of the global enrollment and received at least one dose of randomized study treatment in Cohort A or in Cohort B, will be included in the safety

analysis set. If a patient receives at least one dose of olaparib study treatment they will be summarized in the olaparib arm for safety summaries (e.g. olaparib arm will include patients randomized to olaparib who receive at least one dose of olaparib or those patients randomized to investigator choice arm who receive at least one dose of olaparib study treatment in error at any time). If a patient randomized to olaparib receives only investigator choice of either enzalutamide or abiraterone acetate then they will be summarized as part of the investigator choice arm. Safety data captured on patients receiving investigator choice who have subsequently switched to olaparib upon progression will be summarized per the treatment at the time of the onset of safety condition or lab result and reported in a separate section.

2.1.4 Safety switch analysis set

All patients randomised to investigator choice, who received at least one dose of study treatment in Cohort A or in Cohort B, who have subsequently switched to olaparib upon progression and received at least one dose of olaparib will be included in the safety switch analysis set.

2.1.5 PK analysis set

All patients who have received at least one dose of study medication and provided at least one post-dose analyzable plasma sample for PK analysis will be included in the PK analysis set. Patients with major protocol deviations including changes to the procedures that may impact the quality of the data, or any circumstances that can alter the evaluation of the PK may be excluded from the PK analysis set. These deviations and changes will be specified in a separate protocol deviation specification document.

Table 6 Summary of primary and secondary outcome variables and analysis populations

Outcome Variables	Analysis Populations
Efficacy Data - rPFS (Cohort A, Cohorts A+B, Cohort B) - Time to pain progression (Cohort A, Cohorts A+B, Cohort B) - Overall survival (Cohort A, Cohorts A+B, Cohort B) - Time from randomization to first symptomatic skeletal–related event (Cohort A, Cohort A+B) - Time from randomization to the date of opiate use for cancer-related pain. (Cohort A, Cohorts A+B) - PSA Response (Cohort A, Cohorts A+B) - CTC conversion rate (Cohort A, Cohorts A+B) - PFS2 (Cohort A, Cohorts A+B)	Cohort A FAS, Cohort B FAS, Cohort A+B FAS

Table 6 Summary of primary and secondary outcome variables and analysis populations

Outcome Variables	Analysis Populations
- ORR (Cohort A, Cohorts A+B, Cohort B)	Cohort A EFR, Cohort B EFR, Cohort A+B EFR (Subjects with measurable disease at baseline)
- Duration of Response (Cohort A, Cohorts A+B, Cohort B)	Cohort A EFR, Cohort B EFR, Cohort A+B EFR (Subjects with measurable disease at baseline)
Demography (Cohort A, Cohort A+B, Cohort B)	Cohort A FAS, Cohort B FAS, Cohort A+B FAS
Disposition (Cohort A, Cohorts A+B, Cohort B)	Cohort A FAS, Cohort B FAS, Cohort A+B FAS, Safety switch analysis set
Plasma concentration	PK
HRQoL (Cohort A, Cohorts A+B)	Cohort A FAS, Cohort A+B FAS
Safety data (Cohorts A+B) - Compliance and exposure - Adverse events - Lab measurements - Vital signs	Safety analysis set, Safety switch analysis set
- Concomitant medications	Cohort A FAS, Cohort B FAS, Cohort A+B FAS

2.2 Violations and deviations

The important protocol deviations (IPD) will be listed and summarised by randomized treatment group. None of the deviations will lead to any patients being excluded from the efficacy or safety analysis sets.

The following general categories will be considered IPDs:

- Patients randomized but who did not receive olaparib/investigators choice of NHA.
- Patients who deviate from key entry criteria per the Clinical Study Protocol:

  1) Histologically confirmed diagnosis of prostate cancer.

  2) Candidate for treatment with enzalutamide or abiraterone acetate with documented current evidence of metastatic castration-resistant prostate cancer, where metastatic status is defined as at least one (1) documented metastatic lesion on either bone scan or CT/MRI scan. Subjects whose disease spread is limited to regional pelvic lymph nodes or local recurrence (e.g. bladder, rectum) are not eligible.

3) Subjects must have progressed on prior NHA (e.g. abiraterone acetate and/or enzalutamide) for the treatment of metastatic prostate cancer and/or CRPC. Determination of progression is done per local investigator.

4) Qualifying HRR mutation in tumor tissue by the Lynparza HRR Assay.

- Baseline RECIST or Bone scan > 42 days before start date of randomized treatment, or no baseline RECIST 1.1 or no bone scan assessment on or before date of randomization.
- Persistently missing important protocol required safety assessments (hematology, liver function test, chemistry panel) and potentially having major impact to patient safety (clinical review on a case by case base).
- Received prohibited other anti-cancer agents during study treatment period.
- Met study treatment discontinuation criteria but continued study treatment and potentially had major impact to patients' safety according to clinical judgement.
- Patients randomized who received their randomized study treatment at an incorrect dose or received an alternative study treatment to that which they were randomized.
- Patients assigned to the incorrect cohort.
- Patients who were randomized to investigators choice of NHA and started olaparib treatment before disease progression determined by BICR.

The categorisation of these as IPDs is not automatic and will depend on duration and the perceived effect on efficacy and safety. In addition to the programmatic determination of the deviations above, monitoring notes or summaries will be reviewed to determine any important post entry deviations that are not identifiable via programming, and to check that those identified via programming are correctly classified. The final classification will be made prior to database lock and all decisions will be made whilst blinded to study treatment allocation. For example, details of disallowed concomitant medication use will be reviewed by a physician using blinded data and may be determined as important.

A 'deviation bias' sensitivity analysis will be performed on the rPFS endpoint excluding patients with deviations that may affect the efficacy of the trial therapy if > 10% of patients in either treatment group have IPDs.

The need for such a sensitivity analysis will be determined following review of the protocol deviations ahead of database lock, and will be documented prior to the primary analysis being conducted.

3. PRIMARY AND SECONDARY VARIABLES

The primary assessment of efficacy is rPFS, defined as disease progression according to RECIST 1.1 (for soft tissue disease) and/or PCWG-3 criteria (for bone disease), or death by any cause, whichever comes first. To ensure comparability, identical imaging techniques should be used for the assessment of response at baseline and throughout the study. Further details of the methods used to determine the RECIST response and PCWG3 progression are detailed below and also in Appendix E of the clinical study protocol. The primary analysis will be based on BICR of the radiological scans.

For efficacy analyses, when an event has occurred, every attempt will be made to establish the exact date of the event and enter this into the database. If this is not possible, partial dates will be accepted. If the date of event is not known, then the patient will have an imputed event date as the day of their last known alive event free date prior to DCO.

For the date variables of historical data (i.e., any data referring to the period prior to the informed consent date), if the year is missing then the value will not be imputed. If the month or day is missing, the value will be imputed: month will be imputed with June; day will be imputed as 15th.

3.1 Derivation of RECIST visit responses – malignant soft tissue

For all patients, the RECIST tumor response data will be used to determine each patient's visit response according to RECIST version 1.1. It will also be used to determine if and when a patient has progressed in accordance with RECIST and their best objective response to study treatment.

The baseline assessments of all imaging modalities should be performed as close as possible to the start of study treatment and no more than 4 weeks (-28 days) before randomization.

Following the baseline assessment, subsequent assessments should be performed every 8 weeks (± 7 days), relative to the date of randomization, until objective radiological disease progression by BICR, even after the investigator has deemed objective disease progression, irrespective of treatment decisions or dose interruptions.

If a patient has been deemed to have objective disease progression according to investigator assessment, but not by BICR, he is not eligible to switch to olaparib at that time. Patients should continue to receive randomized study treatment until progression determined by BICR.

If an unscheduled assessment was performed and the subject has not progressed, every attempt should be made to perform the subsequent assessments at their originally scheduled visits. This schedule is to be followed in order to minimize any unintentional bias caused by some subjects being assessed at a different frequency than other subjects.

3.1.1 Target lesions (TLs) – site investigator data

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (LD), (except lymph nodes which must have short axis ≥ 15 mm) with computed tomography (CT) or magnetic resonance imaging (MRI) and which is suitable for accurate repeated measurements. A patient can have a maximum of five measurable lesions recorded at baseline with a maximum of two lesions per organ (representative of all lesions involved and suitable for accurate repeated measurement) and these are referred to as target lesions (TLs). If more than one baseline scan is recorded then measurements from the one that is closest and prior to randomisation will be used to define the baseline sum of TLs. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement. In which circumstance the next largest lesion, which can be measured reproducibly, should be selected.

All other lesions (or sites of disease) not recorded as TL should be identified as non-target lesions (NTLs) at baseline. Measurements are not required for these lesions, but their status should be followed at subsequent visits.

Note: For patients who do not have measurable disease at entry (i.e. no TLs) but have non-measurable disease, evaluation of overall visit responses will be based on the overall NTL assessment and the absence/presence of new lesions (see section 3.1.3 for further details). If a patient does not have measurable disease at baseline then the TL visit response will be not applicable (NA).

Table 7 TL Visit Responses (RECIST 1.1)

Visit Responses	Description
Complete response (CR)	Disappearance of all TLs. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10mm.
Partial response (PR)	At least a 30% decrease in the sum of diameters of TLs, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Progressive disease (PD)	A ≥ 20% increase in the sum of diameters of TLs and an absolute increase of ≥ 5mm, taking as reference the smallest sum of diameters since treatment started including the baseline sum of diameters.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD.

Visit Responses	Description
Not evaluable (NE)	Only relevant in certain situations (i.e. if any of the TLs were not assessed or not evaluable or had a lesion intervention at this visit; and scaling up could not be performed for lesions with interventions). Note: If the sum of diameters meets the progressive disease criteria, progressive disease overrides not evaluable as a TL response.
Not applicable (NA)	No TLs are recorded at baseline.

Rounding of TL data

For calculation of PD and PR for TLs percentage changes from baseline and previous minimum should be rounded to one decimal place before assigning a TL response. For example 19.95% should be rounded to 20.0% but 19.94% should be rounded to 19.9%

Missing TL data

For a visit to be evaluable then all TL measurements should be recorded. However, a visit response of PD should still be assigned if any of the following occurred

- A new lesion is recorded
- A NTL visit response of PD is recorded
- The sum of TLs is sufficiently increased to result in a 20% increase, and an absolute increase of ≥ 5mm, from nadir even assuming the non-recorded TLs have disappeared

Note: the nadir can only be taken from assessments where all the TLs had a LD recorded.

If there is at least one TL measurement missing and a visit response of PD cannot be assigned, the visit response is NE.

Lymph nodes

For lymph nodes, if the size reduces to < 10mm then these are considered non-pathological. However, a size will still be given and this size should still be used to determine the TL visit response as normal. In the special case where all lymph nodes are < 10mm and all other TLs are 0mm then although the sum may be > 0mm the calculation of TL response should be over-written as a CR.

TL visit responses subsequent to CR

CR, PD or NE. can only follow a CR. If a CR has occurred then the following rules at the subsequent visits must be applied:

- Step 1: If all lesions meet the CR criteria (i.e. 0mm or < 10mm for lymph nodes) then response will be set to CR irrespective of whether the criteria for PD of TL is also met i.e. if a lymph node LD increases by 20% but remains < 10mm.

- Step 2: If some lesion measurements are missing but all other lesions meet the CR criteria (i.e. 0mm or < 10mm for lymph nodes) then response will be set to NE irrespective of whether, when referencing the sum of TL diameters, the criteria for PD are also met.

- Step 3: If not all lesions meet the CR criteria (i.e. a pathological lymph node selected as TL has short axis >10mm or the reappearance of previously disappeared lesion) or a new lesion appears then response will be set to PD

- Step 4: If after steps 1 – 3 a response can still not be determined the response will be set to remain as CR

TL too big to measure

If a TL becomes too big to measure this should be indicated in the database and a size ('*x*') above which it cannot be accurately measured should be recorded. If using a value of *x* in the calculation of TL response would not give an overall visit response of PD, then this will be flagged and reviewed by the study team blinded to treatment assignment. It is expected that a visit response of PD will remain in the vast majority of cases.

TL too small to measure

If a TL becomes too small to measure then this will be indicated as such on the case report form and a value of 5mm will be entered into the database and used in TL calculations. However a smaller value may be used if the radiologist has not indicated 'too small to measure' on the case report form and has entered a smaller value that can be reliably measured. If a TL response of PD results then this will be reviewed by the study team blinded to treatment assignment.

Irradiated lesions/lesion intervention

Previously irradiated lesions (i.e. lesion irradiated prior to entry into the study) should be recorded as NTLs and should not form part of the TL assessment.

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation / palliative surgery / embolization), should be handled in the following way and

once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of tumours:

- Step 1: the diameters of the TLs (including the lesions that have had intervention) will be summed and the calculation will be performed in the usual manner. If the visit response is PD, this will remain as a valid response category.

- Step 2: If there was no evidence of progression after step 1, treat the lesion diameter (for those lesions with intervention) as missing and if ≤ 1/3 of the TLs have missing measurements then scale up as described in the 'Scaling' section below. If the scaling results in a visit response of PD then the patient would be assigned a TL response of PD.

- Step 3: If, after both steps, PD has not been assigned, then, if appropriate (i.e. if ≤ 1/3 of the TLs have missing measurements), the scaled sum of diameters calculated in step 2 should be used, and PR or SD then assigned as the visit response. Patients with intervention are evaluable for CR as long as all non-intervened lesions are 0 (or <10mm for lymph nodes) and the lesions that have been subject to intervention have a value of 0 (or <10mm for lymph nodes) recorded. If scaling up is not appropriate due to too few non-missing measurements then the visit response will be set as NE.

At subsequent visits, the above steps will be repeated to determine the TL and overall visit response. When calculating the previous minimum, lesions with intervention should be treated as missing and scaled up (as per step 2 above).

Scaling (applicable only for irradiated lesions/lesion intervention)

If > 1/3 of TL measurements are missing (because of intervention) then the TL response will be NE, unless the sum of diameters of non-missing TL would result in PD (i.e. if using a value of 0 for missing lesions, the sum of diameters has still increased by 20% or more compared to nadir and the sum of TLs has increased by ≥5mm from nadir).

If ≤ 1/3 of the TL measurements are missing (because of intervention) then the results will be scaled up (based on the sizes at the nadir visit to give an estimated sum of diameters) and this will be used in calculations; this is equivalent to comparing the visit sum of diameters of the non-missing lesions to the nadir sum of diameters excluding the lesions with missing measurements.

Example of scaling

Lesion 5 is missing at the follow-up visit; it had a nadir measure of 29.3cm.

The sum of lesions 1-4 at the follow-up is 26 cm. The sum of the corresponding lesions at the nadir visit is 26.8 cm.

Scale up as follows to give an estimated TL sum of 28.4cm:

$$\frac{26}{26.8} \times 29.3 = 28.4cm$$

CR will not be allowed as a TL response for visits where there is missing data. Only PR, SD or PD (or NE) could be assigned as the TL visit response in these cases. However, for visits with ≤1/3 lesion assessments not recorded, the scaled up sum of TLs diameters will be included when defining the nadir value for the assessment of progression.

Lesions that split in two

If a TL splits in two, then the LDs of the split lesions should be summed and reported as the LD for the lesion that split.

Lesions that merge

If two TLs merge, then the LD of the merged lesion should be recorded for one of the TL sizes and the other TL size should be recorded as 0cm.

Change in method of assessment of TLs

CT, MRI and clinical examination are the only methods of assessment that can be used within a trial, with CT and MRI being the preferred methods and clinical examination only used in special cases. If a change in method of assessment occurs, between CT and MRI this will be considered acceptable and no adjustment within the programming is needed.

If a change in method involves clinical examination (e.g. CT changes to clinical examination or vice versa), any affected lesions should be treated as missing.

3.1.2 Non-target lesions (NTLs) and new lesions – site investigator data

At each visit, the investigator should record an overall assessment of the NTL response. This section provides the definitions of the criteria used to determine and record overall response for NTL at the investigational site at each visit.

NTL response will be derived based on the overall assessment of NTLs as follows:

Table 8 NTL Visit Responses

Visit Responses	Description
Complete response (CR)	Disappearance of all NTLs present at baseline with all lymph nodes non-pathological in size (<10 mm short axis).

Visit Responses	Description
Progressive disease (PD)	Unequivocal progression of existing NTLs. Unequivocal progression may be due to an important progression in one lesion only or in several lesions. In all cases, the progression MUST be clinically significant for the physician to consider changing (or stopping) therapy.
Non-CR/Non-PD	Persistence of one or more NTLs with no evidence of progression.
Not evaluable (NE)	Only relevant when one or some of the NTLs were not assessed and, in the investigator's opinion, they are not able to provide an evaluable overall NTL assessment at this visit. Note: For patients without TLs at baseline, this is relevant if any of the NTLs were not assessed at this visit and the progression criteria have not been met.
Not applicable (NA)	Only relevant if there are no NTLs at baseline.

To achieve 'unequivocal progression' on the basis of NTLs, there must be an overall level of substantial worsening in non-target disease such that, even in the presence of SD or PR in TLs, the overall tumour burden has increased sufficiently to merit a determination of disease progression. A modest 'increase' in the size of one or more NTLs is usually not sufficient to qualify for unequivocal progression status.

Details of any new lesions will also be recorded with the date of assessment. The presence of one or more new lesions is assessed as progression.

A lesion identified at a follow up assessment in an anatomical location that was not scanned at baseline is considered a new lesion and will indicate disease progression.

The finding of a new lesion should be unequivocal: i.e. not attributable to differences in scanning technique, change in imaging modality or findings thought to represent something other than tumour.

New lesions will be identified via a Yes/No tick box (excluding bone lesions). The absence and presence of new lesions at each visit should be listed alongside the TL and NTL visit responses.

A new lesion indicates progression so the overall visit response will be PD irrespective of the TL and NTL response.

If the question 'Any new lesions since baseline (excluding bone lesions)' has not been answered with Yes or No and the new lesion details are blank this is not evidence that no new lesions are present, but should not overtly affect the derivation.

Symptomatic progression is not a descriptor for progression of NTLs: it is a reason for stopping study therapy and will not be included in any assessment of NTLs.

Patients with 'symptomatic progression' requiring discontinuation of treatment without objective evidence of disease progression at that time should continue to undergo tumour assessments where possible until objective disease progression is observed.

3.1.3 Overall visit response – site investigator data

Table 9 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 9 Overall visit soft tissue responses

Target lesions	Non-Target lesions	New Lesions	Overall soft tissue response
CR	CR	No	CR
CR	NA	No	CR
NA	CR	No	CR
CR	Non CR/Non PD	No	PR
CR	NE	No	PR
PR	Non PD[a] or NE	No	PR
SD	Non PD[a] or NE	No	SD
NA	Non CR/Non PD	No	SD
NA	NA	No	NED
NE	Non PD[a] or NE	No	NE
NA	NE	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

[a] Non PD = CR or Non CR/Non PD or NA.
CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, NE = not evaluable, NA = not applicable (only relevant if there were no TL and/or NTLs at baseline), NED = No Evidence of Disease (only relevant when there is no TL and NTL from baseline).

3.1.4 Bone Lesion Progression using PCWG3

Bone lesions will be assessed by bone scan and will not be part of the RECIST v1.1 malignant soft tissue assessment. If more than one baseline scan is recorded then measurements from the one that is closest and prior to randomization will be used.

All bone lesions (or sites of disease) should be identified at baseline. Their status should be followed at subsequent visits. At each visit an overall assessment of the bone lesion progression should be recorded by the Investigator.

Progression on a bone scan is identified using PCWG3 as follows:

- At the Week 8 scan:

  Two or more new metastatic bone lesions are observed on the first 8-week scan compared to the baseline assessment. The confirmatory scan, performed at least 6 weeks later and preferably no later than the next scheduled visit for a bone scan (ie, Week 16), must show two or more additional new metastatic bone lesions (for a total of four or more new metastatic bone lesions since the baseline assessment) for progression to be documented.

  Note - The first bone scan completed after baseline will be considered the '8-week scan' regardless if taken at week 8 or at an unscheduled assessment.

- After the Week 8 scan:

  Two or more new metastatic bone lesions are observed compared to the 8-week assessment. The confirmatory scan, performed at least 6 weeks later and preferably at the next scheduled visit for a bone scan, must show the persistence of or an increase in the number of metastatic bone lesions compared to the prior scan for progression to be documented.

The date of progression is the date of the scan that first documents the second lesion.

Table 10 provides the definitions for the visit bone progression status for bone lesions.

Table 10 Bone progression status

Non Progressive Disease (Non-PD)	No evidence of progression, or appearance of one new bone lesion, or non-fulfilment of the progression criteria including new lesions without confirmation of progression.
Progressive Disease (PD)	Bone lesions fulfilling the requirements for at least 2 new lesions and confirmation of progression.
Not Evaluable (NE)	Only relevant if no evaluable follow-up bone scan is available

3.1.5 Blinded Independent Central Review (BICR) with RECIST 1.1 and PCWG3 criteria

A planned BICR of all radiological imaging data will be carried out using RECIST version 1.1 for soft tissue lesions and PCWG3 for bone lesions. All radiological scans for all patients (including those at unscheduled visits, or outside visit windows) will be collected on an ongoing basis and sent to an AstraZeneca appointed Contract Research Organization (CRO) for central analysis. The imaging scans will be reviewed by two independent radiologists using both RECIST 1.1 and PCWG3 and will be adjudicated, if required (i.e. two reviewers' review the scans and adjudication is performed by a separate reviewer in case of a disagreement in a timepoint assessment). For each patient, the BICR will define the overall visit response (i.e. the response obtained overall at each visit by assessing TLs, NTLs and new lesions) data and no programmatic derivation of visit response is necessary. RECIST assessments/scans contributing towards a particular visit may be performed on different dates and for the central review the date of progression for each reviewer will be provided based on the earliest of the scan dates of the component that triggered the progression. The records from the selected reviewer will be used to report all BICR information including dates of progression, visit response, censoring and changes in target lesion dimensions. Endpoints (of ORR, rPFS and DoR) will be derived programmatically from this information.

The independent review charter contains the details of the BICR conducted by the AstraZeneca- appointed Contract Research Organisation (CRO) and has been developed in advance at the start of the study. The BICR will provide RECIST measurements and response and PCWG3 progression status for each visit (i.e. for visits where progression is/is not identified) for each patient at the time of the primary DCO. After the primary rPFS analysis, BICR review of scans will no longer be required.

3.2 Primary endpoint- Radiological Progression Free Survival (rPFS)

The analysis of the primary endpoint, rPFS of cohort A, will be based on tumor assessments determined by BICR using RECIST 1.1 (soft tissue) and PCWG3 (bone) criteria. There is no plan for BICR to read any scans dated after the date of DCO for the primary analysis. There will be no need to request confirmation of BICR PD after this time point, and the investigator-assessed radiographic progression will prevail. [redacted] A sensitivity analysis based on the programmatically derived rPFS based on investigator recorded assessments will be performed.

Radiological progression-free survival is defined as the time from randomization until the date of objective disease progression (soft tissue or bone) or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomized therapy or receives another anti-cancer therapy prior to progression (i.e. date of rPFS event or censoring – date of randomization + 1).

Patients who have not progressed (defined as CR, PR or SD by RECIST 1.1 for soft tissue disease, or Non-PD for bone disease) or died at the time of analysis will be censored at the

time of the earliest date of their last evaluable RECIST 1.1 assessment (taking the latest target lesion, non-target lesion or new lesion scan date) or bone scan assessment that showed Non-PD. Else the latest of the previous RECIST1.1 assessment and bone scan if done at the same visit

However, if the patient progresses or dies immediately after 2 or more consecutive missed visits for either soft tissue or bone assessments, the patient will be censored at the earliest of the previous RECIST 1.1 assessment (taking the latest target lesion, non-target lesion or new lesion scan date) or previous bone scan assessment prior to the two consecutive missed visits (if RECIST and bone scan done at different visits). Else the latest of the previous RECIST1.1 assessment and bone scan if done at the same visit. If the patient has no evaluable visits or does not have baseline data they will be censored at Day 1 unless they die within 2 visits of baseline (in which case their date of death will be used).

With 8 weekly scheduled scans, the allowable interval from the previous radiographic assessment (earliest of the previous RECIST 1.1 assessment or previous bone scan assessment) equates to 18 weeks (126 days), allowing for early and late visits (i.e. 2 x 8 weeks + 1 week for an early assessment + 1 week for a late assessment), or 17 weeks if immediately after the baseline scan (as no need to allow for an early assessment).

The rPFS time will always be derived based on scan dates not visit dates.

When the Investigator is in doubt as to whether PD has occurred and therefore reassesses the patient at a later date, the date of the initial scan should be declared as the date of progression if the repeat scans confirm progression.

CT/MRI and bone scans contributing towards a particular visit may be performed on different dates. The following rules will be applied:

- For BICR (RECIST 1.1 and PCWG3) assessments, the date of progression will be determined based on the earliest of the scan dates of the component that triggered the progression for the adjudicated reviewer selecting PD, or of the reviewer with the earliest date of progression if there is no adjudication for BICR data.

- For investigator assessments, the date of progression will be determined based on the earliest of the dates of the component that triggered the progression

- For BICR and investigator assessments, when censoring a patient for rPFS, the patient will be censored at the earliest of the of the previous RECIST 1.1 assessment (taking the latest target lesion, non-target lesion or new lesion scan date) or previous bone scan assessment.

Table 11 provides the definitions how the visit responses for soft tissue (according to RECIST1.1 criteria) and bone progression status (according to PCWG3 criteria) are combined to give an overall radiological objective visit response.

Table 11 Overall radiological visit response

Overall visit soft tissue response (RECIST 1.1)[a]	Bone progression status (PCWG3)[b]	Bone lesions at visit Present/Absent	Overall radiological visit response
CR	Non-PD	Absent	CR
CR	Non-PD	Present	PR
CR	NE	-	PR
PR	Non-PD or NE	Any	PR
SD	Non-PD or NE	Any	SD
NED	Non-PD	Any	Non-PD
NED	NE	Any	NE
NE	Non-PD or NE	Any	NE
PD	Any	Any	PD
Any	PD	Any	PD

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, NE = not evaluable (if an assessment is missing, it will be considered NE), NED = No Evidence of Disease (only relevant if there were no TL and NTLs at all visits)
[a] See section 3.1.3.
[b] See section 3.1.4.

In order to derive an overall radiological response, the BICR RECIST 1.1 and PCWG3 assessments will be merged by the BICR visit number. The investigator assessments cannot be merged by visit number, they will instead be merged using windows around the protocolled visit schedule as described in the ADaM specification.

3.3 Secondary endpoints

3.3.1 Confirmed Overall Objective Response Rate (ORR)

For patients in the EFR analysis set (who have measurable disease at baseline determined by BICR), objective response rate assessed by BICR (RECIST 1.1 and PCWG3), is defined as the number (%) of patients with at least one visit response of CR or PR, in their soft tissue disease assessed by RECIST 1.1, in the absence of progression on bone scan assessed by PCWG3. For each treatment group, the objective response rate (ORR) is the number of patients with a CR and PR divided by the number of patients in the treatment group .

A confirmed response of CR/PR means that a response of CR/PR is recorded at 1 visit and confirmed by repeat imaging not less than 4 weeks after the visit when the response was first observed with no evidence of progression between the initial and CR/PR confirmation visit. Data obtained up until progression, or last evaluable assessment in the absence of progression, will be included in the assessment of ORR. Patients who discontinue randomized treatment without progression, receive a subsequent anti-cancer therapy (note that for this analysis

radiotherapy is not considered a subsequent anti-cancer therapy) and then respond will not be included as responders in the ORR.

In patients without a confirmed response, an unconfirmed response of CR/PR means that a response of CR/PR is recorded but either no confirmation assessment is performed or a confirmation assessment is performed but response is not confirmed.

In the case where a patient has two non-consecutive visit responses of PR, then, as long as the time between the 2 visits of PR is greater than 4 weeks and there is no PD between the PR visits, the patient will be defined as a confirmed responder. Similarly, if a patient has visit responses of CR, NE, CR, then, as long as the time between the 2 visits of CR is greater than 4 weeks, then a best confirmed response of CR will be assigned.

Overall response rate based on soft tissue will be defined on the basis of RECIST 1.1 only. ORR will also be calculated based on investigator assessment using the EFR analysis set with measurable disease at baseline determined by investigator assessment.

3.3.2 Analgesic Use Scoring

The Analgesic Quantification Algorithm (AQA) developed by Chung et al 2014 will be used to quantify and score analgesic use in the study. The AQA is an eight-point scale that assigns a score as follows:

- 0=No analgesic
- 1=Non-opioid analgesics
- 2=Weak opioids (e.g. codeine, tramadol)
- 3=Strong opioids ≤75 mg oral morphine equivalence (OME) per day
- 4=Strong opioids >75–150 mg OME per day
- 5=Strong opioids >150–300 mg OME per day
- 6=Strong opioids >300–600 mg OME per day
- 7=Strong opioids >600 mg OME per day

Average daily opiate use (based on OME) will be computed using the sum of all opiates used over the 7 days per the assessment schedule. The average daily OME will require at least 4 days of data and will be used to assign the AQA score. An increase of 1 point or more in the AQA score from a starting value of 1 or higher OR ≥ 2 points in AQA score from a starting value of 0 is considered a clinically meaningful increase in opiate use. Similarly, a decrease of 1 point or more in the AQA score from a starting value of 2 or higher is considered a clinically meaningful decrease in opiate use.

3.3.3 Time to Pain Progression (TTPP)

Time to pain progression (based on average BPI-SF worst pain [Item 3] and analgesic [AQA] score) is defined as time from randomization to time point at which worsening in pain (based on BPI-SF worst pain [Item 3]) is observed (i.e. date of pain progression – date of randomization + 1) for asymptomatic patients and symptomatic patients (at baseline) as follows:

Asymptomatic patients at baseline (average BPI-SF worst pain [Item 3] score of 0 and not taking opioids)

- Increase of 2 or more points from baseline in the average BPI-SF worst pain [Item 3] score observed at 2 consecutive follow-up assessments/visits (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit). The date of pain progression will be the earliest date of the assessments contributing to the average of 7-day assessments for BPI-SF [Item 3].

Or

- Initiation of opioid use for pain.

Symptomatic patients at baseline (average BPI-SF Item 3 score >0 and/or currently taking opioids)

- Increase of 2 or more points from baseline in the BPI-SF worst pain [Item 3] score observed at 2 consecutive follow-up assessments/visits (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit) and an average worst pain score ≥4, and no decrease in average opioid use measured as 1 or more points decrease in AQA score from a starting value of 2 or higher. The date of pain progression will be the earliest date of the assessments contributing to the average of 7-day assessments for BPI-SF worst pain [Item 3].

Or

- Increase in the average opioid use measured as 1 or more points increase (or at least 2 points increase if the starting value is 0) in the AQA score from baseline observed at 2 consecutive follow-up assessments/visits (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit).

Information on all analgesics used by patients in pain control will be collected using the analgesic log. For the purposes of pain severity progression, only information on the actual pain medication collected with the analgesic log will be used.

Any BPI-SF worst pain [Item 3] or analgesic log assessments on or before the date of first treatment will be considered a screening assessment.

The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls half way between the two visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between two consecutive visits then the upper limit will be taken as the midpoint value minus 1 day. Study day will be calculated in relation to date of first treatment. For example:

- Day 29, visit window 2 – 42
- Day 57, visit window 43 – 70
- Day 85, visit window 71 – 98

For the Week 4 (Day 29) visit, if there are overlapping screening measurements in the week 4 window on Day 2, 3, 4 etc, resulting in 2 sets of observations in the Week 4 window, then the set of assessments closest to the target day will be used.

Where the average of BPI-SF worst pain [Item 3] score and average AQA score are taken over 7 days for each visit, the 7 day window for both BPI-SF worst pain [Item 3] score and AQA score will start from the date of first entry of the BPI-SF worst pain [Item 3] for that visit. For example, if there are medications entered in the analgesic log prior to the first entry of BPI-SF worst pain [Item 3], the data will not be used in the average AQA score. Additionally, if there are medications entered in the analgesic log after the 7 day period, these will not be used in the average AQA score.

To calculate the average BPI-SF worst pain [Item 3] score over 7 days, there must be at least 4 days with the BPI-SF worst pain [Item 3] completed. The denominator for the average BPI-SF worst pain [Item 3] over 7 days will be the number of days the BPI-SF worst pain [Item 3] is filled in.

To calculate the average AQA score, there must be at least 4 out of the 7 days with evaluable data. To count a day as having evaluable data, at least the BPI-SF worst pain [Item 3] or the analgesic log must be filled in. The denominator for the average AQA score will be the number of days either the BPI-SF worst pain [Item 3] or the analgesic log is filled in.

Pain progression is set to missing at a visit if there are < 4 days data for BPI-SF worst pain [Item 3] and the average AQA score does not meet the progression criteria. If average AQA score meets the progression criteria regardless of available BPI-SF worst pain [Item 3] then the visit is set to progression.

For patients who receive a subsequent anti-cancer therapy, data will only be included until the start date of the subsequent anti-cancer therapy. Note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy. For patients who switch from investigators choice of NHA to olaparib upon progression, olaparib will be considered subsequent therapy.

Patients who do not satisfy the pain progression criteria for asymptomatic patients and symptomatic patients (at baseline) will be censored as follows:

- If a patient meets the criteria for pain progression after 2 or more missed visits (visits which showed < 4 days of BPI-SF worst pain [Item 3] assessments and the average AQA score does not meet the progression criteria), then the patient will be censored at the time of the latest evaluable average BPI-SF worst pain [Item 3] assessment (the earliest date of the assessments contributing to the average will be used).

- Patients who have not met the criteria for pain progression at the time of analysis

  – The censoring date will be the date of the latest evaluable average BPI-SF worst pain [Item 3] assessment (the earliest date of the assessments contributing to the average will be used).

  – Patients with no evaluable baseline or post-baseline data will be censored at Day 1.

- Patients who receive subsequent anti-cancer therapy

  – The censoring date will be the date of the latest evaluable average BPI-SF worst pain [Item 3] assessment prior to the start date of subsequent anti-cancer therapy (the earliest date of the assessments contributing to the average will be used).

  – Patients with no evaluable baseline or post-baseline data will be censored at Day 1.

- Patients who are randomized but do not receive study treatment will be censored at Day 1.

Missing data

Analgesic or pain medication use allows patients to add new medications as "Other" to the handheld device.

In the case where there are reconciled or unreconciled "other" pain medication entries in the analgesic log and the OME value is missing, but the medication is clearly identified as a non-opioid, then the OME value will be set to 0. OME values will not be imputed for reconciled "other" medications which are not clearly identified as non-opioids.

In the case where there are unreconciled "other" pain medication entries which are not clearly identified as non-opioids, OME values will be imputed at two levels while AQA scores will be imputed where OME values cannot be assigned as follows:

Daily completion level:

- If additional pain medications were taken alongside "Other" for a specific day, the highest OME value of the pain medications (based on completed entries) will be selected as the imputed value for each unreconciled "other" entry for the specific day.

7-day completion period level:

- If no additional pain medications were taken alongside "Other" for a specific day, the highest OME value of pain medications (based on completed entries) across the 7 days of assessments will be selected as the imputed value for each unreconciled "other" entry.

AQA score level:

- If no additional pain medications were taken alongside "Other" across the 7 days of assessments, the highest AQA value from all previous visits will be selected as imputed AQA value for the time point.

- If additional pain medications taken alongside "Other" over the 7 days of assessments, and they are all non-opioids, then AQA score of 1 will be assigned (i.e. non-opioid analgesics).

- Where no AQA score can be imputed, the AQA value for that visit will be considered missing.

3.3.4 Brief Pain Inventory – short form (BPI-SF)

The BPI-SF will be used to assess the impact of pain on daily life. The BPI-SF comprises a total of 15 items measuring 2 domains: pain severity and pain interference.

Pain severity subscale/domain

The BPI-SF pain severity domain/subscale consists of 4 items (#3, #4, #5, and #6) that assess pain at its "worst," "least," "average," and "now" (current pain) respectively on an 11-point numeric rating scale (NRS) ranging from 0=No pain to 10=Pain as bad as you can imagine. Pain severity subscale or composite score from all the 4 items will be calculated as a mean score where all items must be non-missing. The average pain severity subscale/domain score at each visit will be calculated as the average of 7 days starting from the date of the first BPI-SF pain severity domain/subscale entry. There must be at least 4 out of the 7 days with a non-missing pain severity subscale score to calculate the average pain severity subscale/domain score.

Time to pain severity progression

Time to 'pain severity' progression will be assessed from the date of randomization (i.e. date of pain severity progression – date of randomization +1) as follows for asymptomatic and symptomatic patients:

Asymptomatic patients at baseline (average "pain severity" subscale score of 0 and not taking opioids)

- Increase of ≥ 2 points in the average "pain severity" subscale score from baseline observed at 2 consecutive follow-up assessments/visits (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit). The date of pain severity progression will be the earliest date of the assessments contributing to the average of 7-day assessments.

Or

- Initiation of any opioid use for pain

Symptomatic patients at baseline (average "pain severity" subscale score >0 and/or currently taking opioids)

- Increase of ≥ 2 points in the average "pain severity" subscale score from baseline observed at 2 consecutive follow-up assessments/visits (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit) and an average "pain severity" subscale score ≥4, and no decrease in average opioid use, measured as 1 or more points decrease in AQA score from a starting value of 2 or higher. The date of pain severity progression will be the earliest date of the assessments contributing to the average of 7-day assessments.

Or

- Increase in the average (i.e. average of 7-day assessments) opioid use measured as 1 or more points increase (or at least 2 points increase if the starting value is 0) in the AQA score from baseline observed at 2 consecutive follow-up assessments/visits (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit). The date of pain severity progression will be the earliest date of the assessments contributing to the average of 7-day assessments.

Information on all analgesics used by patients in pain control will be collected using the analgesic log. For the purposes of pain severity progression, only information on the actual pain medication collected with the analgesic log will be used. For AQA imputation rules for missing data, see section 3.3.3.

Any BPI-SF "pain severity" subscale or analgesic log assessments taken on or before the date of first dose will be considered a screening assessment.

The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls half way between the two visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between two consecutive visits then the upper limit will be taken as the midpoint value minus 1 day. Study day will be calculated in relation to date of first treatment. For example:

- Day 29, visit window 2 – 42
- Day 57, visit window 43 – 70
- Day 85, visit window 71 – 98

For the Week 4 (Day 29) visit, if there are overlapping screening measurements in the week 4 window on Day 2, 3, 4 etc, resulting in 2 sets of observations in the Week 4 window, then the set of assessments closest to the target day will be used.

Where the average of BPI-SF "pain severity" subscale score and average AQA score are taken over 7 days for each visit, the 7 day window for both BPI-SF "pain severity" subscale score and average AQA score will start from the date of first entry of the BPI-SF "pain severity" subscale for that visit. For example, if there are medications entered in the analgesic log prior to the first entry of BPI-SF "pain severity" subscale, the data will not be used in the average AQA score. Additionally, if there are medications entered in the analgesic log after the 7 day period, these will not be used in the average AQA score.

To calculate the average BPI-SF "pain severity" subscale score, there must be at least 4 days with the BPI-SF "pain severity" subscale completed. The denominator for the average BPI-SF "pain severity" subscale will be the number of days the BPI-SF "pain severity" subscale is calculated.

To calculate the average AQA score, there must be at least 4 out of the 7 days with evaluable data. To count a day as having evaluable data, at least the BPI-SF "pain severity" subscale score or the analgesic log must be filled in. The denominator for the average AQA score will be the number of days either the BPI-SF "pain severity" subscale score or the analgesic log is filled in.

Pain severity progression is set to missing at a visit if there are < 4 days data for "pain severity" subscale score and the average AQA score does not meet the progression criteria. If average AQA score meets the progression criteria regardless of available BPI-SF "pain severity" subscale score then the visit is set to progression.

For patients who receive a subsequent anti-cancer therapy data will only be included until the start date of the subsequent anti-cancer therapy. Note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy. For patients who switch from investigators choice of NHA to olaparib upon progression, olaparib will be considered subsequent therapy.

A number of situations will lead to a patient's time to pain severity progression being censored. These are:

- If a patient meets the criteria for pain severity progression after 2 or more missed visits (visits which showed < 4 days of BPI-SF "pain severity" subscale assessments and the average AQA score does not meet the progression criteria), then the patient will be censored at the time of the latest evaluable average BPI-SF "pain severity" subscale assessment (the earliest date of the assessments contributing to the average will be used).

- Patients who have not met the criteria for pain severity progression at the time of analysis
  - The censoring date will be the date of the latest evaluable average BPI-SF "pain severity" subscale assessment (the earliest date of the assessments contributing to the average will be used).
  - Patients with no evaluable baseline or post-baseline data will be censored at Day 1.
- Patients who receive subsequent anti-cancer therapy
  - The censoring date will be the date of the latest evaluable average BPI-SF "pain severity" subscale assessment prior to the start date of subsequent anti-cancer therapy (the earliest date of the assessments contributing to the average will be used).
  - Patients with no evaluable baseline or post-baseline data will be censored at Day 1.
- Patients who are randomized but do not receive study treatment will be censored at Day 1.

Pain interference domain

The BPI-SF pain interference domain includes 7 items: general activity, mood, walking ability, normal work, relations with other people, sleep, and enjoyment of life. The pain interference domain is scored as the mean of the 7 interference items. At least 50% of the items, or 4 out of 7, must have a response for a mean score to be calculated. The average pain interference subscale score at each visit will be calculated as the average of 7 days starting from the date of the first entry. There must be at least 4 out of the 7 days with a non-missing pain interference subscale score to calculate the average pain interference subscale score.

PRO compliance

Summary measures of overall compliance and compliance over time will be derived for BPI-SF. These will be based upon the following definitions:

- Received form: a form that has been received and has a completion date and at least one individual item completed.
- Expected form: a form that is expected to be completed at a scheduled assessment time e.g. a form from a patient who has not withdrawn from the study at the scheduled assessment time but excluding patients in countries with no available translation. For patients that have progressed, the latest of progression and safety follow-up will be used to assess whether the patient is still under BPI-SF follow-up

at the specified assessment time. Date of study discontinuation will be mapped to the nearest visit date to define the number of expected forms. BPI-SF forms are to be completed for 6 months after either progression or treatment discontinuation (whichever comes second).

- Evaluable form: a form with a completion date and at least one subscale that is non-missing.
- Completed questionnaire: a form with all questions completed
- Overall BPI-SF compliance rate is defined as the total number of evaluable forms across all time points, divided by total number of forms expected to be received across all time points multiplied by 100.
- Overall patient compliance rate is defined for each randomized treatment group as the total number of patients with both an evaluable baseline and at least one evaluable follow-up form (as defined above), divided by the total number of patients expected to have completed at least a baseline BPI-SF form multiplied by 100.

Compliance over time will be calculated separately for each visit, including baseline, as the number of patients with an evaluable form at the time point (as defined above), divided by number of patients still expected to complete forms at that visit. Similarly, the evaluability rate over time will be calculated separately for each visit, including baseline, as the number of evaluable forms (per definition above), divided by the number of received forms. Completion rate will be calculated separately for each visit, including baseline, as the number of completed questionnaires (per definition above), divided by the number of received questionnaires. Finally, patient disposition of PRO assessments over time will be computed cumulatively at each visit using tables and bar charts. Descriptive summaries for patient/form disposition will include patients expected to provide PRO assessments and patients unexpected to provide PRO assessment due to death, disease progression and other reasons respectively at each visit.

3.3.5 Pain Palliation

Pain Palliation is defined for patients with an average BPI-SF worst pain [Item 3] score ≥ 4 points at baseline and is assessed as the proportion of patients with a decrease of ≥2 points in BPI-SF worst pain [Item 3] score at 12 weeks, confirmed at least 2 weeks later, without a ≥1 point increase (or ≥2 increase if starting value is 0) in AQA analgesic score.

Note: Confirmation at least 2 weeks later: (start date of subsequent visit – end date of initial visit is >= 14 days).

Information on all analgesics used by patients in pain control will be collected using the analgesic log. For the purposes of pain severity progression, only information on the actual

pain medication collected with the analgesic log will be used. For AQA imputation rules for missing data, see section 3.3.3.

Any BPI-SF worst pain [Item 3] or analgesic log assessments on or before the date of first treatment will be considered a screening assessment.

The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls half way between the two visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between two consecutive visits then the upper limit will be taken as the midpoint value minus 1 day. Study day will be calculated in relation to date of first treatment. For example:

- Day 29, visit window 2 – 42
- Day 57, visit window 43 – 70
- Day 85, visit window 71 – 98

For the Week 4 (Day 29) visit, if there are overlapping screening measurements in the week 4 window on Day 2, 3, 4 etc, resulting in 2 sets of observations in the Week 4 window, then the set of assessments closest to the target day will be used.

Where the average of BPI-SF worst pain [Item 3] score and average AQA score are taken over 7 days for each visit, the 7 day window for both BPI-SF worst pain [Item 3] score and average AQA score will start from the date of first entry of the BPI-SF worst pain [Item 3] for that visit. For example, if there are medications entered in the analgesic log prior to the first entry of BPI-SF worst pain [Item 3], the data will not be used in the average AQA score. Additionally, if there are medications entered in the analgesic log after the 7 day period, these will not be used in the average AQA score.

To calculate the average BPI-SF worst pain [Item 3] score over 7 days, there must be at least 4 days with the BPI-SF worst pain [Item 3] completed. The denominator for the average BPI-SF worst pain [Item 3] over 7 days will be the number of days the BPI-SF worst pain [Item 3] is filled in.

To calculate the average AQA score, there must be at least 4 out of the 7 days with evaluable data. To count a day as having evaluable data, at least the BPI-SF worst pain [Item 3] or the analgesic log must be filled in. The denominator for the average AQA score will be the number of days either the BPI-SF worst pain [Item 3] or the analgesic log is filled in.

Pain palliation is set to missing at a visit if there are < 4 days data for BPI-SF worst pain [Item 3].

For patients who receive a subsequent anti-cancer therapy, data will only be included until the start date of the subsequent anti-cancer therapy. Note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy. For patients who switch from investigators choice of NHA to olaparib upon progression, olaparib will be considered subsequent therapy.

3.3.6 Overall Survival

Overall survival is defined as the time from the date of randomization until death due to any cause regardless of whether the patient withdraws from randomized therapy or receives another anti-cancer therapy (i.e. date of death or censoring – date of randomization + 1). Any patient not known to have died at the time of analysis will be censored based on the last recorded date on which the patient was known to be alive. This variable is recorded within the survival status module of the eCRF (SUR_DAT, recorded within the SURVIVE module of the eCRF).

For overall survival, if a partial date exists when an event has occurred then the following imputation method will be used: if month is missing, then impute month with January. If the day is missing, then impute day with 01. Then choose the latest date from the imputed event date and the patient's last known alive date+1 as the final imputed event date.

Note: Survival calls will be made in the week following the date of DCO for the primary rPFS analysis, final OS analysis [redacted], and if patients are confirmed to be alive or if the death date is post the DCO date these patients will be censored at the date of DCO.

3.3.7 Time to first Symptomatic Skeletal –Related Event (SSRE)

Time from randomization to first symptomatic skeletal–related event as defined by any of the following or a combination:

- Use of radiation therapy to prevent or relieve skeletal symptoms
- Occurrence of new symptomatic pathological bone fractures (vertebral or nonvertebral). Radiologic documentation is required.
  A pathological fracture, as determined by investigator, is defined as associated with low or no trauma and deemed to have occurred at a site of bone metastasis
- Occurrence of spinal cord compression. Radiologic documentation required
- Orthopedic surgical intervention for bone metastasis

Patients who have not experienced any of the above conditions will be censored at time of death, or time of last SSRE assessment.

3.3.8 Duration of Response (DoR)

For patients in the EFR analysis set (who have measurable disease at baseline determined by BICR) and have a confirmed response (CR or PR as described in section 3.3.1), duration of response (DoR) will be defined as the time from the date of first documented confirmed response until date of documented progression or death in the absence of disease progression (i.e. date of rPFS event or censoring – date of first confirmed response + 1).

The time of the first confirmed response will be defined as the latest of the dates contributing towards the first visit response of confirmed PR or CR. The end of response will be defined as the date of progression or death from any cause used for the rPFS endpoint. If a patient does not progress following a response, then their duration of response will use the rPFS censoring date as the date at which that patient is censored for DoR. However, if the date of rPFS censoring is on or before the date of the first confirmed response then the patient will be censored at Day 1 for DoR.

The time to response is the time from randomization to the first onset of a confirmed objective tumor response (i.e. date of first confirmed response – date of randomization + 1).

An unconfirmed response is defined in section 3.3.1. Duration of response will also be calculated for patients with an unconfirmed response for BICR and investigator assessments separately.

3.3.9 Time to Opiate Use for Cancer Pain

Time to Opiate use is defined as the time from randomization to the date of opiate use for cancer-related pain on subjects who have not received any opiates at baseline. Subjects who have not received opiates during the study or died prior to receiving opiates will be censored at the last study assessment date prior to DCO where no opiate use was recorded.

3.3.10 Prostate Specific Antigen (PSA) Response

PSA response is defined as the proportion of patients achieving a ≥50% decrease in PSA from baseline to the lowest post-baseline PSA result, confirmed by a second consecutive PSA assessment at least 3 weeks later. For patients who receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy), data will only be included until the start date of the subsequent anti-cancer therapy. All patients in the FAS will be included, regardless of having a baseline PSA measurement.

- A patient will be regarded as having a single PSA visit response if their PSA level at any post-dose visit is reduced by 50% or more compared with baseline

- A patient will be regarded as having a confirmed PSA response if they have a reduction in PSA level of 50% or more compared with baseline that is confirmed at the next assessment at least 3 weeks later (i.e., decrease relative to baseline of at least 50% documented on 2 consecutive occasions at least 3 weeks apart).

For the calculation of PSA responses, values of the form of "< x" (i.e. below the lower limit of quantification) or > x (i.e. above the upper limit of quantification) will be imputed as "x".

3.3.11 PSA changes on continuous scale

PSA changes on a continuous scale will be evaluated in patients in the FAS who have a valid baseline and post-baseline PSA measurement. Patients without a baseline PSA measurement and/or a post-baseline PSA measurement will be excluded from the analysis. For patients who

receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy), data will only be included until the start date of the subsequent anti-cancer therapy.

- PSA levels will be evaluated in terms of percentage change from baseline which will be derived for each post baseline visit where PSA data are available:

- Percentage change from baseline = [(post-dose PSA level – baseline PSA level) / baseline PSA level] *100

- Best percentage change from baseline in PSA will be derived as the biggest reduction in PSA level compared with baseline (or the smallest increase in the absence of a reduction) taking account of all PSA values collected for each patient.

For the calculation of PSA changes on a continuous scale, values of the form of "< x" (i.e. below the lower limit of quantification) or > x (i.e. above the upper limit of quantification) will be imputed as "x".

3.3.12 Circulating Tumor Cell (CTC) Conversion rate

Defined as the proportion of patients achieving a decline in the number of CTCs from ≥5 cells/7.5 mL at baseline to <5 cells/7.5 mL at any visit post baseline. For patients who receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy), data will only be included until the start date of the subsequent anti-cancer therapy. All patients in the FAS will be included for CTC conversion rate, regardless of having a baseline CTC measurement.

For the calculation of CTC conversion rates, CTC values of the form of "< x" (i.e. below the lower limit of quantification) or > x (i.e. above the upper limit of quantification) will be imputed as "x".

CTC counts on a continuous scale will be evaluated in patients in the FAS who have a valid baseline and post-baseline CTC measurement. For patients who receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy), data will only be included until the start date of the subsequent anti-cancer therapy. Patients without a baseline CTC measurement and/or a post-baseline CTC measurement will be excluded from the analysis.

3.3.13 Time from randomization to second progression or death (PFS2)

Defined as the time from the date of randomization to the earliest of the investigator assessed progression event (subsequent to that used for the primary variable rPFS) or death (i.e. date of PFS2 event or censoring – date of randomisation + 1). The date of second progression will be recorded by the investigator and defined according to local standard clinical practice and may involve any of; objective radiological, symptomatic progression or death. The date of the PFS2 assessment and investigator opinion of progression status (progressed or non-progressed) at each assessment will be recorded in the electronic case report form (eCRF).

Patients alive and for whom a second disease progression has not been observed should be censored at the last time known to be alive and without a second disease progression (i.e., censored at the latest of the rPFS date by investigator assessment or PFS2 assessment date if the patient has not had a second progression or death). However, if the patient experiences a second progression or dies after two or more visits where there was no evaluable PFS2 assessment (i.e. the evaluable PFS2 assessment was greater than 196 days since the prior evaluable assessment, based on two 12-weekly visits plus two allowed 2 week visit windows) the patient will be censored at the time of the prior evaluable PFS2 assessment.

3.3.14 Functional Assessment of Cancer Therapy- Prostate Cancer (FACT-P)

The patient-reported FACT-P will be used to assess health-related quality of life. The questionnaire will be administered, at baseline, Week 8, 16 and 24 and then continue to be administered to all patients (who have not withdrawn consent) every 8 weeks.

The following outcome measures will be calculated from the FACT-P questionnaire, the resulting value is the total score for the associated questions or scaled scores:

- Physical well-being subscale (PWB) (Questions GP1 to GP7)
- Social/family well-being subscale (SWB) (Questions GS1 to GS7)
- Emotional well-being subscale (EWB) (Questions GE1 to GE6)
- Functional well-being subscale (FWB) (Questions GF1 to GF7)
- Prostate cancer subscale (PCS) (Questions C2, C6, P1 to P8, BL2 and BL5)
- Total Functional Assessment of Cancer Therapy- General (FACT-G) score, sum of PWB, SWB, EWB and FWB
- Trial Outcome Index (TOI), sum of PWB, FWB and PCS
- Functional Assessment of Prostate Cancer Symptoms Index 6 (FAPSI-6) (Questions P1 to P3, GP1, C2 and GE6)
- Total FACT-P score (sum of scores of all the sub-scales)

Items to be reversed

- Each question in the FACT-P questionnaires has a choice of 5 responses, “Not at all”, “A little bit”, “Somewhat”, “Quite a bit” and “Very much”. The scores range from 0 (“Not at all”) to 4 (“Very much”) for positively phrased questions. Negatively phrased questions have a reverse scoring, from 0 (“Very much”) to 4 (“Not at all”). This results in a consistent approach, where higher scores indicate a better quality of life.

- Note, questions that are reversed (via subtraction of the response from 4) are: GP1-7, GE1, GE3-6, C2, P1-3, P6-P8 and BL2.

Missing data

As per the functional assessment of chronic illness (FACIT) scoring guidelines (Cella et al 1993, Cella et al 1994, Esper et al 1997),

- More than 80% of questions in a questionnaire must be completed for the questionnaire to have the FACT-P total score evaluable. If 80% or less of questions are completed, the FACT-P total scores will not be calculated. Similarly, FACT-G total score and and TOI score require more than 80% of the relevant questions to be completed for the score to be evaluable.

- For each domain (PWB, SWB, EWB, FWB and PCS) if more than 50% of the items were answered (e.g., a minimum of 4 of 7 items, 4 of 6 items, etc), the subscale score will be calculated by multiplying the sum of subscale by the number of items in the subscale, then dividing by the number of items actually answered:

  Subscale score= (sum of item scores x N of items in subscale)/ N of items answered

- If at least 50% of the domain items are missing, that domain will be treated as missing and thus NE. The total score for each variable (FACT-G, FACT-P TOI and FACT-P Total) is then calculated as the sum of the un-weighted prorated scores. If a domain score is NE, any health related quality of life (HRQL) variable which these domains contribute to is also termed NE. For example, for the FACT-P TOI variable, if PWB is NE at a visit, the FACT-P TOI variable is also NE at this visit. Also, the FACT-P total score cannot be computed if any of the domain scores is NE.

Visit responses

The last non-missing assessment before randomization will be assigned to be the baseline assessment.

At each post-baseline visit, the following criteria as listed below in Table 12 will be used to assign a visit response for the FACT-P total score, TOI, FACT-G, FAPSI-6, PCS, PWB and FWB scores (Cella et al 2009). This response should be maintained for 2 consecutive visits.

Table 12 Definition of visit response for FACT-P, FAPSI-6, FACT-G, TOI, PCS, FWB and PWB

FACT-P scale	Change from baseline	Visit response
FACT-P-Total	≥ +10	Improved
	≤ -10	Worsened

Table 12 Definition of visit response for FACT-P, FAPSI-6, FACT-G, TOI, PCS, FWB and PWB

FACT-P scale	Change from baseline	Visit response
	Otherwise (i.e. >-10 and <+10)	No change
	Missing/non-calculable score	Not evaluable
FAPSI-6	≥ +3	Improved
	≤ -3	Worsened
	Otherwise (i.e. >-3 and <+3)	No change
	Missing/non-calculable score	Not evaluable
TOI	≥ +9	Improved
	≤ -9	Worsened
	Otherwise (i.e. >-9 and <+9)	No change
	Missing/non-calculable score	Not evaluable
FACT-G	≥ + 7	Improved
	≤ - 7	Worsened
	Otherwise (i.e. >-7 and <+7)	No change
	Missing/non-calculable score	Not evaluable
PCS, FWB, PWB	≥ +3	Improved
	≤ -3	Worsened
	Otherwise (i.e. >-3 and <+3)	No change
	Missing/non-calculable score	Not evaluable

Note for some patients it will not be immediately possible to obtain a visit response for a particular subscale, for example:

- Patients with no baseline score for a particular subscale, or no baseline data at all
- Patients whose baseline subscale score is too close to the maximum or minimum possible score to allow an increase or decrease of the specific size to be observed.
  - For patients whose baseline score is greater than the maximum possible score for that subscale minus the score needed to satisfy improvement, the best visit response possible will be “No Change”.

- For patients whose baseline score is less than the threshold needed for worsening (e.g., a baseline FACT-P TOI < 9) all post-baseline visit responses will be considered not-calculable.

For those patients who meet the criteria above (where it is not possible to improve or worsen), descriptive data will be provided.

At the conclusion of the study, the criteria listed in Table 13 will be used to assign a best overall response score based on the individual visit responses. For patients who receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy), data will only be included until the start date of the subsequent anti-cancer therapy.

Table 13 Overall score response criteria

Best overall response	Criteria
Improved	Two consecutive visit responses of 'improved'. Consecutive visits need to be at least 3 weeks apart.
No Change	Does not qualify for overall score response of 'improved'. Two consecutive visit responses of either 'no change', or 'improved' and 'no change'
Worsened	Does not qualify for overall score response of 'improved' or 'no change'. A visit response of 'worsened'
Not evaluable	Missing or non-calculable scores
Other	Does not qualify for one of the above

Time to deterioration for FACT-P

Time to deterioration in HRQL as measured by FACT-P total score will be defined as the interval from the date of randomization until the date of the first clinically meaningful deterioration that is confirmed at a subsequent visit at least 3 weeks apart with no improvement in between the visits (except if it was the patient's last available assessment) or death (by any cause) in the absence of a clinically meaningful deterioration, regardless of whether the patient discontinues study drug(s) prior to the deterioration in FACT-P total score. Death will be included as an event only if it occurs within 2 PRO assessment visits from the last available PRO assessment. For patients who receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy), data will only be included until the start date of the subsequent anti-cancer therapy. Time to deterioration as measured by FACT-P TOI, FACT-G, FAPSI-6, PCS, PWB and FWB will be derived similarly.

A worsening is as described in Table 12, for example, for FACT-P TOI a decrease in score from baseline of ≥ 9 will constitute a deterioration. Improvement is also as defined in Table 12.

Radiologic progression will not be considered as deterioration in symptoms.

Note, under the same principles applied to the primary outcome variable (rPFS), time to deterioration will be derived regardless of whether the patient withdraws from randomised therapy prior to symptom deterioration. A number of situations will lead to a patient's time to deterioration of HRQL endpoints being censored. These are:

- If a patient either dies or meets the criteria for deterioration after 2 or more missed HRQL assessments, then the patient will be censored at the time of the latest evaluable HRQL assessment. These patients will be presented as e.g. "Censored FACT-P total score" in summaries.

- Patients who have not met the criteria for symptom deterioration or died at the time of analysis will be censored at the time of the latest evaluable HRQL assessment:

  - The censoring date will be the date of the last assessment that led to evaluable being assigned for FACT-P total score. These patients will be presented as alive and deterioration-free in summaries.

  - Patients with no evaluable baseline or post-baseline data will be censored at Day 1 unless they die within 2 visits of baseline (in which case their date of death will be used). These patients will be presented as censored in summaries.

- Patients who receive subsequent anti-cancer therapy

  - The censoring date will be the date of the last assessment prior to the start date of subsequent anti-cancer therapy that led to evaluable being assigned for FACT-P total score. These patients will be presented as alive and deterioration-free in summaries.

  - Patients with no evaluable baseline or post-baseline data will be censored at Day 1 unless they die within 2 visits of baseline (in which case their date of death will be used). These patients will be presented as censored in summaries.

- Patients whose baseline subscale score is close to the minimum possible

  - For patients whose baseline score is less than the threshold needed for worsening (e.g., a baseline FACT-P total score of < 6), time to deterioration will be censored at Day 1 unless they die within 2 visits of baseline. Patients who haven't died will be presented as "Censored FACT-P Total Score" in summaries.

The time to deterioration of HRQL will be derived based on assessment dates, not visit dates.

PRO compliance

Summary measures of overall compliance and compliance over time will be derived for the FACT-P. These will be based upon the compliance derivations described for BPI-SF.

3.3.15 Pharmacokinetic Endpoint

Olaparib plasma concentrations will be measured at Week 4 (Visit 3) pre-dose (- 30 min ± 15 min) and at 30 min ± 15 min, 2 ± 0.5 hour, 5 ± 0.5 hour, and 8± 1 hour post-dose.

3.4 Exploratory endpoints

3.4.1 PRO-CTCAE

The PRO-CTCAE consists of nominal categories (e.g. “none” to “very severe” for some items in the questionnaire) it will be collected every 2 weeks (starting day 1) for first 8 weeks and every 4 weeks thereafter until 6 months after discontinuation of study treatment.

PRO Compliance

Summary measures of overall compliance and compliance over time will be derived for the PRO-CTCAE. These will be based upon the compliance derivations described for BPI-SF.

3.4.2 Patient Global Impression of Change (PGIC)

The PGIC item is included to assess how a patient perceives their overall change in health status since the start of study treatment. Patients will choose from response options ranging from “Very Much Improved” to “Very Much Worse.” This item is useful in characterizing the overall impact of the treatment. The PGIC assessments will be performed on week 8 and every 8 weeks thereafter.

PRO Compliance

Summary measures of overall compliance and compliance over time will be derived for the PGIC. These will be based upon the compliance derivations described for BPI-SF.

3.4.3 Health Economics

3.4.3.1 Resource use

To investigate the impact of treatment and disease on health care resource, the following variables will be captured:

- Planned and unplanned hospital attendances beyond trial protocol mandated visits (including physician visits, emergency room visits, day cases and admissions)
- Primary sign or symptom the patient presents with
- Length of hospital stay

- Length of any time spent in an intensive care unit (ICU)

The study site staff will complete the "Hospital Admission (HOSPAD)" eCRF at the site at every scheduled clinic visit up to and including the study treatment discontinuation follow up visit. If the subject discontinues study treatment for reasons other than RECIST progression, the "HOSPAD" eCRF should continue to be administered until progression has been determined by BICR.

Where admitted overnight, the length of hospital stay will be calculated as the difference between the date of hospital discharge (or death date) and the start date of hospitalisation or start of study drug if the start of study drug is after start date of hospitalisation (length of hospital stay = end date of hospitalisation – start date of hospitalisation + 1). Patients with missing discharge dates will be calculated as the difference between the last day with available data and the start date of hospitalisation. The length of ICU stay will be calculated using the same method.

Further Payer required analyses involving resource use will be described in detail in the Payer analysis plan.

3.4.3.2 EQ-5D-5L

The EQ-5D-5L index comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). For each dimension, respondents select which statement best describes their health on that day from a possible 5 options of increasing levels of severity (no problems, slight problems, moderate problems, severe problems and unable to/ extreme problems). A unique EQ-5D health state is referred to by a 5-digit code allowing for a total of 3125 health states. For example, state 11111 indicates no problems on any of the 5 dimensions. These data will be converted into a weighted health state index by applying scores from EQ-5D value sets elicited from general population samples (the base case will be the UK valuation set, with other country value sets applied in scenario analyses). Where values sets are not available, the EQ-5D-5L to EQ-5D-3L crosswalk will be applied. In addition to the descriptive system, respondents also assess their health today on a visual analogue scale (VAS), ranging from 0 (worst imaginable health) to 100 (best imaginable health).

Measurements are collected on study day 1 and every 8 weeks thereafter until 24 weeks following radiographic progression.

Further analyses involving each of the 5 dimensions of health, the EQ-5D-VAS and utilities will be needed to support Payer dossiers and will be described in detail in the payer analysis plan.

PRO Compliance

Summary measures of overall compliance and compliance over time will be derived for the EQ-5D-5L. These will be based upon the compliance derivations described for BPI-SF.

3.5 Safety

3.5.1 General considerations for safety assessments

Baseline will generally be the last value obtained prior to the first dose of study medication. Alternatively, if two visits are equally eligible to assess patient status at baseline (e.g., screening and baseline assessments both on the same date prior to first dose with no washout or other intervention in the screening period), the average can be taken as a baseline value. For non-numeric laboratory tests (i.e. some of the urinalysis parameters) where taking an average is not possible then the best value would be taken as baseline as this is the most conservative. In the scenario where there are two assessments on day 1, one with time recorded and the other without time recorded, the one with time recorded would be selected as baseline.

Time windows will be defined for any presentations that summarise values by visit. The following conventions will apply:

- The time windows will be exhaustive so that data recorded at any time point has the potential to be summarised. Inclusion within the time window will be based on the actual date and not the intended date of the visit.

- All unscheduled visit data have the potential to be included in the summaries.

- The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls half way between the two visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between two consecutive visits then the upper limit will be taken as the midpoint value minus 1 day. For example, the visit windows for laboratory assessment data (with 4 weeks between scheduled assessments) are:

    - Day 29, visit window 2 – 42
    - Day 57, visit window 43 – 70
    - Day 85, visit window 71 – 98

- For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment will be used (regardless of where it falls in an interval).

- Listings should display all values contributing to a time point for a patient.

- For visit based summaries

    - If there is more than one value per patient within a time window then the closest value to the scheduled visit date will be summarised, or the earlier, in the event the values are equidistant from the nominal visit date. The listings will highlight the value for the patient that contributed to the summary table, wherever feasible. Note: in summaries of extreme values all post baseline

values collected are used including those collected at unscheduled visits regardless of whether or not the value is closest to the scheduled visit date.

- To prevent very large tables or plots being produced that contain many cells with meaningless data, for each treatment group, visit data should only be summarised if the number of observations is greater than the minimum of 20 and > 1/3 of patients dosed.

- For summaries at a patient level, all values will be included, regardless of whether they appear in a corresponding visit based summary, when deriving a patient level statistic such as a maximum.

Where safety data are summarised over time, study day will be calculated in relation to date of first treatment.

Missing safety data will generally not be imputed. However, safety assessment values of the form of "< x" (i.e. below the lower limit of quantification) or > x (i.e. above the upper limit of quantification) will be imputed as "x" in the calculation of summary statistics but displayed as "< x" or "> x" in the listings. Additionally, adverse events that have missing causality (after data querying) will be assumed to be related to study drug.

For laboratory data the following applies:

- Numerical summaries will provide the mean, standard deviation, median, minimum, maximum, and lower and upper quartile for visit based tabular summaries.

3.5.2 Handling of partial dates

For missing start dates for AEs and concomitant medications/procedures, the following will be applied:

- Missing day - Impute the 1st of the month unless month is the same as month of first dose of study drug then impute first dose date.

- Missing day and month – impute 1st January unless year is the same as first dose date then impute first dose date.

- Completely missing date – impute first dose date unless the end date is less than the first dose date, in which case impute the 1st January of the same year as the end date.

When imputing a start date ensure that the new imputed date is sensible e.g. is prior to the end date of the AE.

For missing stop dates of AEs or concomitant medications/procedures, the following will be applied:

- Missing day - Impute the last day of the month unless month is the same as month of the last dose of study drug then impute last dose date.

- Missing day and month – impute 31st December unless year is the same as last dose date then impute last dose date.

- Completely missing date – do not impute.

The imputation of dates will be used to decide if an observation is treatment emergent for adverse events or concomitant medications. The imputed dates are not used to calculate durations. Where partial dates occur, listings will contain the date collected in the partial form.

3.5.3 Adverse Events

The definitions of adverse events (AEs) and serious AEs (SAEs) are given in Sections 6.1 and 6.2 of the clinical study protocol. AEs and SAEs will be collected throughout the study, from date of informed consent until 30 days after the last dose of study treatment (or end of follow-up period). Events will be defined as treatment emergent if they onset, or worsen (by investigator report of a change in intensity), on or after the first dose date, and up to and including 30 days following the date of last dose of study medication. The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) will be used to code the AEs. AEs will be graded according to the National Cancer Institute Common Terminology Criteria for AEs (CTCAE Version 4.03).

Other significant Adverse Events (OAE)

During the evaluation of the AE data, an AstraZeneca medically qualified expert will review the list of AEs that were not reported as SAEs or discontinuations of IP due to AEs (DAEs). Based on the expert's judgement, significant AEs of particular clinical importance may, after consultation with the Global Patient Safety Physician, be considered other significant AEs (OAEs) and reported as such in the Clinical Study Report. A similar review of laboratory/vital signs/ECG data will be performed for identification of OAEs.

AEs of special interest

Adverse events of special interest (AESI) are events of scientific and medical interest specific to the further understanding of olaparib's safety profile and require close monitoring and rapid communication by the investigators to AstraZeneca. An AESI may be serious or non-serious.

Adverse events of special interest for olaparib are:

- Important Potential Risks of MDS/AML

- New primary malignancy (other than MDS/AML)

- Pneumonitis

3.5.4 Concomitant medications

Concomitant medications will be classified according to the current version of the WHO Drug Dictionary.

Concomitant medications will be classed as either:

1. Concomitant medications starting prior to first dose (pre-study)

2. Concomitant medications starting on or after first dose date (on study). Medications that start on the same day as the first dose of study treatment will be assumed to occur after study treatment has been administered, and be classified as on-study.

3.5.5 Laboratory assessments

Blood samples for determination of clinical chemistry, hematology and coagulation will be taken at each scheduled visit and urine samples to determine urinalysis will be taken at screening and Day 1 visits. The laboratory parameters to be collected are given in Section 5.2.1 of the protocol.

3.5.6 Exposure

Study drug exposure (days) for olaparib will be defined as time from first dose of olaparib, up to and including the, last day of dosing of olaparib. Exposure to investigators choice of enzalutamide or abiraterone acetate will be calculated in the same way using enzalutamide or abiraterone acetate only. Exposure to prednisone/prednisolone will not be calculated.

Exposure will be defined as:

Last dose date – first dose date + 1.

Exposure to olaparib for patients randomised to investigator choice of treatment, who have subsequently switched to olaparib upon progression will be defined as:

Last dose date of olaparib – first dose date of olaparib + 1.

If the last dose date is unknown, the soonest available date afterwards where it is confirmed that no drug is being taken will be used instead.

Actual exposure of olaparib/investigators choice of NHA:

- Actual exposure = intended exposure – total duration of dose interruptions, where intended exposure will be calculated as above and a dose interruption is defined as any length of time where the patient has not taken any of the planned daily dose.

Missed or forgotten doses

Missed and forgotten doses should be recorded on the EX, EX1 and EX2 modules for olaparib, enzalutamide and abiraterone respectively as drug interrupted with the reason recorded as “Subject forgot to take dose”. These missed or forgotten doses will not be included as dose interruptions in the summary tables but the information will appear in the listing for dosing. However, these missed and forgotten doses will be considered in the derivation of actual exposure.

Safety Follow-up

Total Safety Follow-up = min((last dose date + 30 days), date of withdrawal of consent, date of death, date of DCO, date of first dose of subsequent anti-cancer therapy) – first dose date +1.

3.5.7 Dose intensity

Relative dose intensity (RDI) is the percentage of the actual dose delivered relative to the intended dose through to treatment discontinuation. RDI will be defined as follows:

- RDI = 100% * d/D, where d is the actual cumulative dose delivered up to the actual last day of dosing and D is the intended cumulative dose up to the or the actual last day of dosing. D is the total dose that would be delivered, if there were no modification to dose or schedule.

Percentage intended dose (PID) is the percentage of the actual dose delivered relative to the intended dose through to progression. PID will be defined as follows:

- PID = 100% * d/D, where d is the actual cumulative dose delivered up to progression (or a censoring event) and D is the intended cumulative dose up to progression (or a censoring event). D is the total dose that would be delivered, if there were no modification to dose or schedule.

Intensity of olaparib, enzalutamide, and abiraterone acetate will be summarised separately. The intended cumulative dose is defined as 300mg olaparib twice daily, 160mg enzalutamide once daily and 1000mg abiraterone acetate once daily.

3.5.8 Vital signs

Vital signs, including BP (mmHg), pulse rate (beats/minute), body temperature (°C) and weight (kg), will be assessed at screening, baseline and as clinically indicated and will be summarized at baseline. Changes in vital signs should be recorded as an AE, if applicable.

3.5.9 Physical examination

Physical examination assessments will be performed at screening, day 1 and as clinically indicated.

3.5.10 ECG

Resting 12-lead ECGs will be performed within 7 days prior to starting study treatment and when clinically indicated. Measurements should be taken after the patient has been rested in a supine position for at least 5 minutes. All ECGs will be assessed locally to determine whether they are clinically significantly abnormal / not clinically significantly abnormal. If there is a clinically significant abnormal finding, it will be recorded as an AE by the Investigator.

4. ANALYSIS METHODS

4.1 General principles

The DCO date for the statistical analysis for the primary objective of the study will be when approximately 143 rPFS events in Cohort A, are expected to have occurred.

This study is comparing olaparib to investigators choice of NHA. Results of statistical analyses will be presented using corresponding 2-sided 95% confidence intervals and 2-sided p-values, where appropriate.

Efficacy data will be summarised and analysed on the full analysis set. Safety data, including exposure data, will be summarised and analysed on the safety population. Study population and demography data will be summarised on the full analysis set.

Descriptive statistics will be used for all variables, as appropriate. Continuous variables will be summarised by the number of observations, mean, standard deviation, median, upper and lower quartiles, minimum, and maximum. Categorical variables will be summarised by frequency counts and percentages for each category. Unless otherwise stated, percentages will be calculated out of the population total and for each treatment group and will be rounded to 1 decimal place.

For continuous data, the mean and median will be rounded to 1 additional decimal place compared to the original data. The standard deviation will be rounded to 2 additional decimal places compared to the original data. Minimum and maximum will be displayed with the same accuracy as the original data.

All analyses will be performed in SAS version 9.2 or later.

In general, for efficacy endpoints the last non missing measurement prior to randomization will be considered the baseline measurement. However, if an evaluable assessment is only available after randomization but before the first dose of randomized treatment then this assessment will be used as baseline. For safety and PRO endpoints, the last observation before the first dose of study treatment will be considered the baseline measurement unless otherwise specified. For assessments on the day of first dose where time is not captured, a nominal pre-dose indicator, if available, will serve as sufficient evidence that the assessment occurred prior to first dose. If neither time nor a nominal pre-dose indicator are present assessments will be considered pre-dose if such procedures are required by the protocol to be conducted before first dose.

In all summaries change from baseline variables will be calculated as the post-treatment value minus the value at baseline. The percentage change from baseline will be calculated as (postbaseline value - baseline value) / baseline value x 100.

4.2 Analysis methods

Table 14 Pre-planned statistical and sensitivity analyses to be conducted

Endpoints analysed	Cohort	Notes
Radiologic progression-free survival	Cohort A Cohort B Cohort A+B	• Stratified log-rank test: -Primary analysis (based on BICR [RECIST 1.1 and PCWG3] assessments and stratified in accordance with the pooling strategy defined in section 4.2.2 • Hazard ratio using Cox proportional hazards model (with ties=Efron and the stratification variables determined by the pooling strategy as covariates) • Plots and summaries of number (%) patients with progression or death events using Kaplan-Meier (KM) method. • Stratified log tank test and cox proportional hazards model will be repeated for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients. • KM plot will be produced for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients.
rPFS sensitivity analysis: (a) Evaluation-time bias (b) Attrition bias (c) Censoring bias (d) Ascertainment bias (e) Sensitivity analysis using unequivocal clinical progression in addition to radiological progression (f) Sensitivity analysis for confirmation of bone progression (g) Sensitivity analysis censoring patients with subsequent therapy or discontinuation of study drug	Cohort A Cohort B Cohort A+B	• Stratified log-rank test stratified in accordance with the pooling strategy (all sensitivity analyses except for censoring bias) • Hazard ratio using Cox proportional hazards model with ties=Efron and the stratification variables determined by the pooling strategy as covariates (all sensitivity analyses except for censoring bias) • KM plot (censoring bias and ascertainment bias only)

Endpoints analysed	Cohort	Notes

rPFS subgroup analysis (based on BICR assessments):
• Previous taxane use (yes, no)
• Measurable disease at baseline (yes, no)
• Metastases at baseline: bone only vs visceral vs other
• ECOG performance status at baseline (0, 1, or 2)
• Age at randomisation (<65, ≥65)
• Region (Asia, Europe, North and South America)
• Race (White, Black/African-American, Asian, Other)
• Baseline PSA (above/below median baseline PSA of the patients across both treatment groups)
	Cohort A
	Cohort A+B
		• HRs and associated 2-sided CIs will be estimated using a Cox proportional hazards model (with the Efron method being used for handling ties) that contains the treatment term, factor and treatment-by-factor interaction term
		• If there are less than 5 events across both treatment arms in a subgroup then descriptive statistics will be provided instead.
		• KM plots for the following subgroups:
		  • Previous taxane use (yes, no)
		  • Measurable disease at baseline (yes, no)
		  • Metastases at baseline: bone only vs visceral vs other
		  • Baseline PSA (above/below median baseline PSA of the patients across both treatment groups)

rPFS subgroup analysis (based on BICR assessments): HRR gene mutations in the full analysis set and confirmed FMI F1CDx patients: • *BRCA1* and/or *BRCA2* • *BRCA1* and/or *BRCA2* and/or *ATM* • *BRCA1* and/or *BRCA2* and/or *ATM* and/or *CDK12* • *BARD1* and/or *BRIP1* and/or *CHEK1* and/or *CHEK2* and/or *FANCL* and/or *PALB2* and/or *PPP2R2A* and/or *RAD51B* and/or *RAD51C* and/or *RAD51D* and/or *RAD54L* • *BRCA1* and/or *BRCA2* and/or *CDK12* • Any single HRR mutation • Each individual gene in patients with single HRR gene mutations only HRR gene mutations in confirmed myriad *gBRCA*m patients: • *BRCA1* only • *BRCA2* only	Cohort A+B	• HRs and associated 2-sided CIs will be estimated using a Cox proportional hazards model (with the Efron method being used for handling ties) that contains the treatment term, factor and treatment-by-factor interaction term • If there are less than 5 events across both treatment groups in a subgroup then descriptive statistics will be provided instead. • KM plots for all combinations of HRR gene mutations in the full analysis set only.

Endpoints analysed	Cohort	Notes
Confirmed objective response rate • BICR assessment using RECIST and bone scan data • BICR assessment using RECIST soft tissue only • Investigator assessment using RECIST and bone scan data • Investigator assessment using RECIST soft tissue only	Cohort A Cohort B Cohort A+B	• Odds ratio using logistic regression adjusted for prior taxane. If there are not at least 5 responses across both treatment groups, then a Fisher's exact test using mid p-values will be used. • Logistic regression for confirmed ORR using BICR assessment (RECIST and bone scan data) will be repeated for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients.
Confirmed objective response rate subgroup analysis: • Each individual gene in patients with single HRR gene mutations only	Cohort A+B	• Odds ratio using logistic regression adjusting for treatment, factor and treatment by factor interaction
Unconfirmed objective response rate • BICR assessment using RECIST and bone scan data • BICR assessment using RECIST soft tissue only • Investigator assessment using RECIST and bone scan data • Investigator assessment using RECIST soft tissue only	Cohort A Cohort B Cohort A+B	• Odds ratio using logistic regression adjusted for prior taxane. If there are not at least 5 responses across both treatment groups then a Fisher's exact test using mid p-values will be used
Time to pain progression: • All patients in FAS • Patients in FAS who have not taken any analgesics at baseline	Cohort A Cohort B Cohort A+B	• Stratified log rank test stratified in accordance with the pooling strategy • Hazard ratio using a Cox proportional hazards model (with ties=Efron and the stratification variables determined by the pooling strategy as covariates) • Plots and summaries of number (%) patients with events using KM method. • Stratified log tank test and cox proportional hazards model will be repeated for confirmed FMI F1CDx patients and

Endpoints analysed	Cohort	Notes
		confirmed myriad *gBRCA*m patients.
Proportion of patients with pain progression: • All patients in FAS • Patients in FAS who have not taken any analgesics at baseline	Cohort A Cohort B Cohort A+B	• Odds ratio using logistic regression adjusted for the stratification variables determined by the pooling strategy
Overall survival	Cohort A Cohort B Cohort A+B	• Stratified log rank test stratified in accordance with the pooling strategy • Hazard ratio using a Cox proportional hazards model (with ties=Efron and the stratification variables determined by the pooling strategy as covariates) • Plots and summaries of number (%) patients with events using KM method. • Stratified log tank test and cox proportional hazards model will be repeated for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients.
Time to first Symptomatic Skeletal-Related Event	Cohort A Cohort A+B	• Stratified log-rank test stratified in accordance with the pooling strategy • Hazard ratio using a Cox proportional hazards model (with ties=Efron and the stratification variables determined by the pooling strategy as covariates) • Plots and summaries of number (%) patients with events using KM method.
Duration of Response: • Confirmed response • Unconfirmed response	Cohort A Cohort B Cohort A+B	• Summarized using descriptive statistics • KM plots
Time to Opiate use for Cancer related Pain	Cohort A Cohort A+B	• Stratified log-rank test stratified in accordance with the pooling strategy • Hazard ratio using a Cox proportional hazards model (with ties=Efron and the stratification variables determined by the

Endpoints analysed	Cohort	Notes
		pooling strategy as covariates) • Plots and summaries of number (%) patients with events using KM method.

Prostate Specific Antigen (PSA) Response	Cohort A
	Cohort A+B
		• Summarized using descriptive statistics
		• Waterfall plots
		  • Best percentage change from baseline
		  • Percentage change from baseline at Week 12
		• Confirmed PSA best response presented with 95% CIs

Circulating Tumor Cell (CTC) conversion rate	Cohort A
	Cohort A+B
		• Summarized using descriptive statistics
		• Waterfall plots
		  • Best change from baseline
		  • Best percentage change from baseline
		• Proportion of patients achieving CTC conversion at any time presented with 95% CIs

Time from randomization to second progression or death	Cohort A Cohort A+B	• Stratified log-rank test stratified in accordance with the pooling strategy • Hazard ratio using a Cox proportional hazards model (with ties=Efron and the stratification variables determined by the pooling strategy as covariates) • Plots and summaries of number (%) patients with events using KM method.

Endpoints analysed	Cohort	Notes
Time to deterioration in FACT-P (FACT-P total score, FACT-G total score, TOI, FAPSI-6, FWB, PWB, PCS)	Cohort A Cohort A+B	• Stratified log rank test stratified in accordance with the pooling strategy • Hazard ratio using a Cox proportional hazards model (with ties=Efron and the stratification variables determined by the pooling strategy as covariates) • Forest plot
FACT-P (FACT-P total score, FACT-G total score, TOI, FAPSI-6, FWB, PWB, PCS)	Cohort A Cohort A+B	• Summary statistics by treatment group • Change from baseline using a MMRM which includes treatment, visit and treatment by visit interaction as explanatory variables and the baseline FACT-P total score as a covariate, along with the baseline FACT-P total score by visit interaction and the stratification variables prior taxane and measurable disease as determined by the pooling strategy
FACT-P improvement rate (FACT-P total score, FACT-G total score, TOI, FAPSI-6, FWB, PWB, PCS)		• Odds ratio using logistic regression adjusted for the stratification variables determined by the pooling strategy. If there are not at least 5 responses across both treatment groups then a Fisher's exact test using mid p-values will be used.
BPI-SF (worst pain [Item 3], pain severity domain and pain interference domain)	Cohort A Cohort A+B	• Change from baseline using a MMRM which includes treatment, visit and treatment by visit interaction as explanatory variables and the baseline FACT-P total score as a covariate, along with the baseline FACT-P total score by visit interaction and the stratification variables prior taxane and measurable disease as determined by the pooling strategy

Endpoints analysed	Cohort	Notes
Time to pain severity • All patients in FAS • Patients in FAS who have not taken any analgesics at baseline	Cohort A Cohort A+B	• Stratified log rank test stratified in accordance with the pooling strategy • Hazard ratio using a Cox proportional hazards model (with ties=Efron and the stratification variables determined by the pooling strategy as covariates) • Plots and summaries of number (%) patients with events using KM method
Proportion of patients with pain severity • All patients in FAS • Patients in FAS who have not taken any analgesics at baseline	Cohort A Cohort A+B	• Odds ratio using logistic regression adjusted for the stratification variables determined by the pooling strategy
Proportion of patients with pain palliation • All patients in FAS • Patients in FAS who have not taken any analgesics at baseline	Cohort A Cohort A+B	• Odds ratio using logistic regression adjusted for the stratification variables determined by the pooling strategy

4.2.1 Multiplicity

The multiple testing procedure (MTP) (as shown in Figure 2) will define which significance levels should be applied to the interpretation of the raw p-values for the primary endpoint of rPFS and the key secondary endpoints.

Hypotheses will be tested using a MTP with an alpha-exhaustive recycling strategy (Burman et al 2009). Upon achieving statistical significance on the primary endpoint rPFS in Cohort A, testing of each of the secondary endpoints, ORR (Cohort A), rPFS (Cohort A+B), time to pain progression (Cohort A), and overall survival (Cohort A) will be performed sequentially with the 2-sided 5% level of alpha recycled from the primary rPFS (Cohort A) endpoint. This testing procedure stops when the entire test mass is allocated to non-rejected hypotheses. Implementation of this pre-defined ordered testing procedure, including recycling, will strongly control type I error at 5% (2-sided), among all key hypotheses.

Note that two analyses of OS are planned:

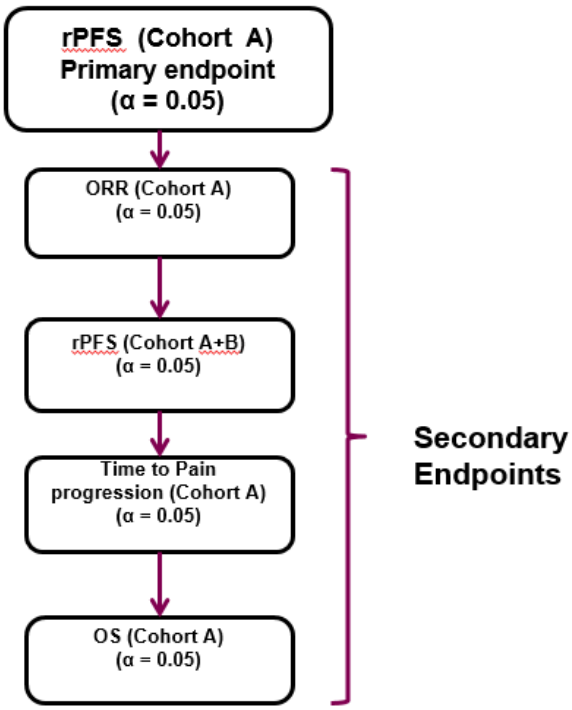
1. An initial look at the time of the primary rPFS analysis, and

2. A final analysis at approximately 61% maturity.

Using an O'Brien-Fleming spending function, the initial look will use approximately 0.012 one sided alpha level with 80% information fraction and a final OS analysis will use as alpha level of 0.021 with approximately 146 events (61% maturity) estimated to occur

approximately 48 months after first patient randomized in the study. The actual information fraction will be calculated based on the observed number of OS events at the time of the initial look and the this will determine the one-side alpha level for the analysis.

Figure 2 Multiplicity strategy maintaining overall Type I error rate



4.2.2 Analysis of the primary efficacy variable (rPFS)

The primary analysis of radiologic progression-free survival will be performed when approximately 143 rPFS events (60% maturity) in cohort A have occurred based on BICR (RECIST 1.1 and PCWG3) assessment.

The primary analysis will be based on the BICR assessment of rPFS using all scans regardless of whether they were scheduled or not. The hypothesis of superiority of olaparib compared to investigator choice will be tested using a log rank test with the Breslow method for handling ties, stratified by the variables determined by the pooling strategy described below.

The effect of olaparib versus investigator choice of NHA will be estimated by the hazard ratio and corresponding 95% confidence interval. This analysis will be performed using a Cox Proportional Hazards Model with the Efron approach being used for handling ties and the stratification variables determined by the pooling strategy being used as covariates. The 2-sided 95% confidence intervals will be calculated using the profile likelihood method and a HR less than 1 will favour olaparib.

Any patients mis-stratified in the IVRS/IWRS will be included in the stratified log rank test using the baseline data collected in the IVRS/IWRS.

Although it is expected that there will be enough rPFS events in each strata (where strata are defined as categories formed from – prior taxane * measurable disease * treatment) to allow a meaningful analysis, if any stratum for either treatment arm contains less than 5 events, then a pooling strategy will be employed. The order of preference for pooling will be (prior taxane * treatment), (measurable disease * treatment), unstratified. In addition, for analyses on Cohort A+B, Cohort will be added as a stratification factor provided that the addition does not lead to <5 events in any strata. Prior taxane and measurable disease will use data collected via IVRS. The pooling strategy will be employed for Cohort A, Cohort B and Cohort A+B separately. All sensitivity analyses and secondary endpoints (except for ORR which only includes prior taxane) will use the same strata as the primary model, for that endpoint, unless there are <5 events per stratum and then an unadjusted model will be used.

Kaplan-Meier (KM) survival curves (product-limit estimates) of rPFS will be presented by treatment group, together with a summary of associated statistics (median rPFS time, and 6- and 12-month survival rate estimates). Summaries of the number and percentage of patients experiencing an rPFS event, and the type of event (RECIST progression, PCWG-3 progression, both or death) will also be presented.

The assumption of proportionality will be assessed. Note that in the presence of non-proportionality, the HR will be interpreted as an average HR over the observed extent of follow-up. Proportionality will be tested firstly by producing plots of complementary log-log (event times) versus log (time) and, if these raise concerns, a time dependent covariate would be fitted (adding a treatment-by-time or treatment-by-ln(time) interaction term) to assess the extent to which this represents random variation. If qualitative non-proportionality is observed then using stratification rather than covariate adjustment will be considered.

Additionally, the rPFS endpoint will be analyzed in Cohort B and Cohort A+B as part of secondary analyses. The rPFS analysis, along with Kaplan Meier curves, will also be produced for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients in Cohort A, Cohort B and Cohort A+B.

4.2.2.1 Subgroup analysis

Subgroup analyses will be conducted for rPFS endpoint. The purpose of the subgroup analyses is to assess the consistency of treatment effect across potential or expected prognostic factors. If there are too few responders or events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less than 5 events across both treatment groups per subgroup), the relationship between that subgroup and the endpoint will not be formally analyzed. In this case, only descriptive summaries will be provided.

The following subgroups of the full analysis set in Cohort A and Cohort A+B will be analyzed for rPFS for stratification factors

- Previous taxane use (yes, no)
- Measurable disease at baseline (yes, no)

Values collected on the eCRF will be used to define subgroups for stratification factors.

Additional subgroups of interest include:

- HRR gene mutations in the FAS , each individual gene, and pre-specified combinations:
    - *BRCA1* and/or *BRCA2* - Cohort A+B
    - *BRCA1* and/or *BRCA2* and/or *ATM* - Cohort A+B
    - *BRCA1* and/or *BRCA2* and/or *ATM* and/or *CDK12* - Cohort A+B
    - *BARD1* and/or *BRIP1* and/or *CHEK1* and/or *CHEK2* and/or *FANCL* and/or *PALB2* and/or *PPP2R2A* and/or *RAD51B* and/or *RAD51C* and/or *RAD51D* and/or *RAD54L* - Cohort A+B
    - *BRCA1* and/or *BRCA2* and/or *CDK12* - Cohort A+B
    - Any single HRR mutation* - Cohort A+B
    - Each individual gene in patients with single HRR gene mutations only - Cohort A+B

    * This will include all patients who only have a single HRR gene mutation in the FAS. There will be two levels for this subgroup, patients with a single HRR gene mutation and patients who do not have a single HRR gene mutation i.e. patients with co-mutations.

    There will be two levels per subgroup (or gene for each individual gene category): yes and no (patients with specified HRR gene mutations and patients not confirmed to have specified gene mutations respectively).

- Metastases at baseline: bone only vs visceral vs other - Cohort A and Cohort A+B
- ECOG performance status at baseline (0, 1, or 2) - Cohort A and Cohort A+B
- Age at randomisation (<65, ≥65) - Cohort A and Cohort A+B
- Region (Asia, Europe, North and South America) - Cohort A and Cohort A+B
- Race (White, Black/African-American, Asian, Other) - Cohort A and Cohort A+B
- Baseline PSA (above/below median baseline PSA of the patients across both treatment groups) - Cohort A and Cohort A+B

The following subgroups will be analysed for confirmed FMI F1CDx patients:

- *BRCA1* and/or *BRCA2* - Cohort A+B
- *BRCA1* and/or *BRCA2* and/or *ATM* - Cohort A+B
- *BRCA1* and/or *BRCA2* and/or *ATM* and/or *CDK12* - Cohort A+B
- *BARD1* and/or *BRIP1* and/or *CHEK1* and/or *CHEK2* and/or *FANCL* and/or *PALB2* and/or *PPP2R2A* and/or *RAD51B* and/or *RAD51C* and/or *RAD51D* and/or *RAD54L* - Cohort A+B
- *BRCA1* and/or *BRCA2* and/or *CDK12* - Cohort A+B
- Any single HRR mutation* - Cohort A+B
- Each individual gene in patients with single HRR gene mutations only - Cohort A+B

* This will include all patients who only have a single HRR gene mutation for confirmed FMI F1CDx patients. There will be two levels for this subgroup, patients with a single HRR gene mutation and patients who do not have a single HRR gene mutation i.e. patients with co-mutations.

There will be two levels per subgroup (or gene for each individual gene category): yes and no (patients with specified HRR gene mutations and patients not confirmed to have specified gene mutations respectively).

The following subgroups will be analysed for confirmed myriad *gBRCA*m patients in patients:

- *BRCA1* only
- *BRCA2* only

There will be two levels per gene: yes and no (patients with specified HRR gene mutations and patients not confirmed to have specified gene mutations respectively).

In each subgroup, the HRs for radiological progression by BICR (olaparib vs investigator choice of NHA) and associated 2-sided CIs will be estimated using a Cox proportional hazards model (with the Efron method being used for handling ties) that contains the treatment term, factor and treatment-by-factor interaction term. For all subgroups except the HRR gene mutation subgroups, the treatment effect HRs for each treatment comparison along with their confidence intervals will be obtained for each level of the subgroup from this single model. For the HRR gene mutation subgroups, the treatment effect HRs for each treatment comparison along with their confidence intervals will only be displayed for the subgroup level containing patients who have the specified HRR gene mutations. The HRs and 95% CIs will be presented on a forest plot including the HR and 95% CI from the overall population (using the primary analysis). No adjustment to the significance level for testing of subgroups will be made since all these subgroup analyses will be considered exploratory as supportive of the primary analysis of rPFS.

In addition, KM plots will be produced for the following subgroups in the full analysis set:

- Previous taxane use (yes, no) Cohort A and Cohort A+B

- Measurable disease at baseline (yes, no) Cohort A and Cohort A+B
- HRR gene mutations in the full analysis set, each individual gene, and pre-specified combinations:
    - *BRCA1* and/or *BRCA2* - Cohort A+B
    - *BRCA1* and/or *BRCA2* and/or *ATM* - Cohort A+B
    - *BRCA1* and/or *BRCA2* and/or *ATM* and/or *CDK12* - Cohort A+B
    - *BARD1* and/or *BRIP1* and/or *CHEK1* and/or *CHEK2* and/or *FANCL* and/or *PALB2* and/or *PPP2R2A* and/or *RAD51B* and/or *RAD51C* and/or *RAD51D* and/or *RAD54L* - Cohort A+B
    - *BRCA1* and/or *BRCA2* and/or *CDK12* - Cohort A+B
    - Any single HRR mutation - Cohort A+B
    - Each individual gene in patients with single HRR gene mutations only - Cohort A+B
- Metastases at baseline: bone only vs visceral vs other - Cohort A and Cohort A+B
- Baseline PSA (above/below median baseline PSA of the patients across both treatment groups) - Cohort A and Cohort A+B

Consistency of treatment effect between subgroups

The presence of quantitative interactions will be assessed by means of an overall global interaction test. This will be performed in the overall population by comparing the fit of a Cox proportional hazards model including treatment, all covariates (stratification factors), and all covariate-by-treatment interaction terms, with one that excludes the interaction terms. This will not include the HRR mutation subgroups. This will be assessed at the 2-sided 10% significance level. If the fit of the model is not significantly improved then it will be concluded that overall the treatment effect is consistent across the subgroups.

If the global interaction test is found to be statistically significant, an attempt to determine the cause and type of interaction will be made. Stepwise backwards selection will be performed on the saturated model, whereby (using a 10% level throughout) the least significant interaction terms are removed one-by-one and any newly significant interactions re-included until a final model is reached where all included interactions are significant and all excluded interactions are non-significant. All main effects will be included in the model regardless of whether the corresponding interaction term is still present. This approach will identify the factors that independently alter the treatment effect and prevent identification of multiple correlated interactions. The p-values reported will represent those from the final model resulting from stepwise backwards selection; the 'selection model'.

Any quantitative interactions identified using this procedure will then be tested to rule out any qualitative interaction using the approach of Gail and Simon (Gail and Simon 1985).

4.2.2.2 Sensitivity analysis

Sensitivity analyses may be performed to assess the possible presence of time-assessment bias (i.e., differential assessment times between treatment groups). Summary statistics for the number of weeks between rPFS time and the last evaluable assessment prior to progression will be presented for each treatment group. For all sensitivity analyses, the same methodology and model will be used as per the primary rPFS analysis, including stratification factors in accordance with the final pooling strategy. The HR and associated 95% CI will be reported. Median rPFS will be presented by treatment group.

The following sensitivity analyses will be evaluated:

(a) Evaluation-time bias

Sensitivity analyses will be performed to assess possible evaluation-time bias that may be introduced if scans are not performed at the protocol-scheduled time points. The midpoint between the time of progression and the previous evaluable assessment (RECIST or PCWG3) will be analyzed as described for the primary analysis of rPFS. The previous evaluable assessment will be the latest of the previous RECIST 1.1 assessment or previous bone scan assessment. Note that midpoint values resulting in non-integer values should be rounded down. For patients whose death was treated as a rPFS event, the date of death will be used to derive the rPFS time used in the analysis. This approach has been shown to be robust to even highly asymmetric schedules (Sun and Chen 2010).

(b) Attrition bias

Attrition bias will be assessed by repeating the primary rPFS analysis except that the actual rPFS event times, rather than the censored time, of patients who progressed or died in the absence of progression immediately following 2, or more, non-evaluable tumor assessments will be included. In addition, patients who take subsequent therapy prior to progression or death will be censored at their last evaluable assessment prior to taking the subsequent therapy, where the last evaluable assessment is the latest of the previous RECIST 1.1 assessment or previous bone scan assessment.

(c) Censoring bias

A KM plot of the time to censoring will be produced where the censoring indicator of the primary rPFS analysis is reversed.

(d) Ascertainment bias

Analysis of rPFS will be based on investigator assessment. A KM plot will also be produced for investigator assessment.

(e) Sensitivity analysis using unequivocal clinical progression in addition to radiological progression

Repeating primary rPFS analysis based on BICR assessed RECIST and bone scan data with the addition of unequivocal progression as an event. Where unequivocal clinical progression is defined as, cancer pain requiring initiation of opioids, need to initiate cytotoxic chemotherapy, radiation therapy or surgical intervention for complications due to tumor progression or deterioration in ECOG performance to >= Grade 3. Unequivocal clinical progression will be determined using the dose discontinuation module (DOSDISC) where patients discontinue due to "Unequivocal clinical progression".

(f) Sensitivity analysis for confirmation of bone progression

Repeat primary rPFS analysis based on BICR assessed RECIST and bone scan data with revised confirmation criteria for bone progression where bone progression accompanied by unequivocal clinical progression does not require a confirmatory bone scan. Unequivocal clinical progression will be determined using the dose discontinuation module (DOSDISC) where patients discontinue due to "Unequivocal clinical progression".

(g) Sensitivity analysis censoring patients with subsequent therapy or discontinuation of study drug

Repeat primary rPFS analysis censoring patients with subsequent therapy or discontinuation of study drug prior to progression (censoring at the earliest date of the first day of subsequent therapy or the last day of study drug).

As a key sensitivity analysis to the primary endpoint of rPFS, the primary analysis will be repeated excluding any patients who did not have a qualifying gene mutation confirmed positive by the FMI F1CDx or the Myriad *gBRCA*m test (where these tests have been performed to support companion diagnostic development), see section 4.2.2. Similarly, the analysis will be repeated in all the 15 HRR qualifying gene mutations (Cohort A+B), see section 4.2.2.1 for specification of these outputs.

4.2.3 Analysis of secondary variables

Analyses of secondary endpoints will be performed at the time of the primary rPFS analysis including an interim analysis for OS. The final analysis of OS will occur upon achieving approximately 146 deaths (61% maturity) in Cohort A.

All time to event analyses in Cohort A, B and A+B will be conducted in accordance with the final pooling strategy for stratification factors in the primary analysis of rPFS, where less than 5 events for a time to event endpoint within each stratum, will result in collapsing of strata until the minimum 5 event criterion is achieved. Unstratified analyses will be conducted for any secondary endpoints that still do not conform to the 5 event rule per stratum. This will also be supported by unstratified sensitivity analyses of the primary endpoint. Additional sensitivity analyses may also be conducted as required.

4.2.3.1 Confirmed Objective response rate (ORR)

Objective response rate will be assessed based on BICR assessed RECIST and bone scan data (using all scans regardless of whether they were scheduled or not) in patients in the EFR analysis set (patients with measurable disease at baseline determined by BICR) .

The ORR will be compared between olaparib and investigator choice using a logistic regression model adjusting for previous taxane (yes, no) collected via IVRS. The results of the analysis will be presented in terms of an odds ratio, with an odds ratio greater than 1 favouring olaparib, together with the associated 95% profile likelihood CI and p-value (based on twice the change in log-likelihood resulting from the addition of a treatment factor to the model). A statistically significant difference in ORR will be demonstrated using a 2-sided 5% alpha level based on the multiplicity strategy described in section 4.2.1.

If there are not at least 5 responses across both treatment groups then a Fisher's exact test using mid p-values will be presented. The mid-p-value modification of the Fisher exact test amounts to subtracting half of the probability of the observed table from Fisher's p-value.

Fisher's exact test mid p-value = Two sided p-value – (Table probability ÷ 2)

Summaries of number (%) of patients with a tumor response (BICR and investigator assessment of CR or PR) will be presented.

Best objective response (BOR) will be assessed based on BICR assessed RECIST and bone scan data in patients in the EFR analysis set (patients with measurable disease at baseline determined by BICR) in the full analysis set. For each treatment group BOR will be summarised by n (%) of patients for each category (CR, PR, SD, PD and NE), no formal statistical analyses for BOR are planned. For patients with a BOR of SD, the number and percentage of patients who had an unconfirmed response of CR/PR will be displayed. BOR will be repeated based on soft tissue response only using RECIST 1.1.

Analysis of ORR will be performed in Cohorts A, B and A+B for the EFR analysis set. Additionally, analysis of soft tissue ORR will be performed in Cohorts A, and A+B for the EFR analysis set. All ORR analyses will be repeated for patients with an unconfirmed response of CR/PR.

All analyses will be repeated using investigator assessed response for patients in the EFR analysis set (patients with measurable disease at baseline determined by investigator assessment).

The confirmed ORR logistic regression will also be produced for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients.

Subgroup analysis

The analysis of confirmed ORR will be repeated in Cohort A+B EFR set using the logistic regression methodology above adjusting for treatment, factor and treatment by factor interaction for the following subgroup:

- Each individual gene in patients with single HRR gene mutations only

There will be two levels per gene, yes and no (patients with specified HRR gene mutation and patients not confirmed to have specified gene mutation respectively). The results of the analysis will be presented in terms of an odds ratio, with an odds ratio greater than 1 favouring olaparib, together with the associated 95% profile likelihood CI. No adjustment to the significance level for testing of subgroups will be made since all these subgroup analyses will be considered exploratory as supportive of the confirmed ORR analysis.

If there are not at least 5 responses across both treatment groups for a subgroup, then the relationship between that subgroup and the endpoint will not be formally analyzed. In this case, only descriptive summaries will be provided.

4.2.3.2 Time to pain progression (TTPP)

Time to pain progression will be analyzed at the time of the primary rPFS analysis using the same methods as in the analysis of rPFS. The HR and 95% confidence interval will be based on the cox proportional hazards model and the p-value on the stratified log rank test stratified in accordance with the pooling strategy described in section 4.2.2 with previous taxane and measurable disease as stratification variables. A 2-sided 5% alpha level will be used to test time to pain progression based on the multiplicity strategy.

TTPP will also be assessed in the subgroup of patients who have not taken any analgesics at baseline. In addition, TTPP will be produced for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients.

A KM plot of time to pain progression will be presented by treatment group. Summaries of the number and percentage of patients experiencing pain progression will be provided along with median time to pain progression for each treatment arm.

The proportion of patients with pain progression will be compared between olaparib and investigator choice using logistic regression adjusting for the stratification variables determined by the pooling strategy. The results of the analysis will be presented in terms of an odds ratio, with an odds ratio less than 1 favouring olaparib, together with the associated 95% profile likelihood CI and p-value (based on twice the change in log-likelihood resulting from the addition of a treatment factor to the model). If there are not at least 5 responses across both treatment groups then a Fisher's exact test using mid p-values will be presented. The mid-p-value modification of the Fisher exact test amounts to subtracting half of the probability of the observed table from Fisher's p-value.

Fisher's exact test mid p-value = Two sided p-value – (Table probability ÷ 2)

This will be repeated patients who are non-opiate users at baseline.

Analysis of Pain progression will be performed in Cohorts A, B and A+B.

Missing data

Tabular summaries will be presented to show the percentage of randomized patients with one or more reconciled or unreconciled "Other" pain medication as well as all imputed OME values by assessment level for each visit for the full analysis set. Percentage of patients with imputed AQA values as applicable will be included in the summary tables.

The sensitivity and robustness of the imputation approaches will be assessed by producing a listing of all patients with one or more imputed OME values. The listing will include:

i) Highest OME value of pre-selected pain medications from Master List

ii) Highest OME value of reconciled "Other" pain medication

iii) Highest imputed OME value of unreconciled "Other" pain medication.

4.2.3.3 Overall survival

Analysis of the secondary efficacy endpoint overall survival will be performed at the time of the primary analysis of rPFS with approximately 117 (49%) events in Cohort A expected at this point in time and will be considered an interim OS analysis. As per the multiplicity strategy, testing of the OS endpoint will utilize the alpha level recycled from rPFS primary endpoint and the secondary endpoints ORR (Cohort A), rPFS (Cohort A+B), and time to pain progression (Cohort A) using a 2-sided 5% alpha spend.

The p-value will be based on the stratified log rank test stratified in accordance with the pooling strategy described in section 4.2.2 . HR and 95% CI will be based on the Cox model.

A KM plot of OS will be presented by treatment group. Summaries of the number and percentage of deaths and those alive and censored will be provided along with median time to death for each treatment arm.

An interim overall survival analysis will take place at the time of the primary rPFS analysis, and a final OS analysis will happen when approximately 146 (61%) of events have occurred. Note that if OS is not tested at the initial look then no formal testing of OS will occur at the final analysis of OS.

Exploratory analyses of OS in Cohort A, adjusting for impact of subsequent PARP inhibitor trial or treatment (or other potentially active investigational agents), may be performed if a sufficient proportion of patients switch. Methods such as Rank Preserving Structural Failure Time (RPSFT) (Robins et al 1991), Inverse Probability of Censoring Weighting (IPCW) (Robins 1993) and other methods in development may be explored. The decision to adjust and final choice of methods will be based on a blinded review of the data and the plausibility of the underlying assumptions.

Overall survival analysis will be performed in Cohorts A, B, and A+B. Overall survival analysis (not including the KM plot) will be repeated for the subset of patients positive by F1CDx and myriad *gBRCA*m. The subgroup analyses specified for rPFS will be repeated for OS for the final analysis only.

4.2.3.4 Time to first Symptomatic Skeletal-Related Event (SSRE)

Time to SSRE will be analyzed using the same methods as in the analysis of the primary endpoint rPFS.

A KM plot of time to SSRE will be presented by treatment group. Summaries of the number and percentage of patients with symptomatic skeletal related events and those who are censored will be provided along with median time to symptomatic skeletal related events for each treatment arm for Cohort A and Cohort A+B only.

4.2.3.5 Duration of Response

Duration of response will be assessed in patients in the EFR analysis set (patients with measurable disease at baseline determined by BICR) for Cohorts A, B and A+B.

Descriptive data will be provided for the duration of response in responding subjects, including the associated KM curves (without any formal comparison of or p-value attached). Descriptive data for onset of response will also be provided. DoR will be repeated for patients with an unconfirmed response.

4.2.3.6 Time to Opiate use for Cancer related Pain

Time to opiate use will be analyzed at the time of the primary rPFS analysis using the same methods as in the analysis of rPFS. The, p-value will be based on the stratified log rank test stratified in accordance with the pooling strategy described in section 4.2.2 HR and 95% CI will be based on the Cox model. A 2-sided 5% alpha level will be used to test time to opiate use based on multiplicity strategy.

A KM plot of time to opiate use will be presented by treatment group. Summaries of the number and percentage of subjects using opiates will be provided along with median time to opiate use for each treatment arm.

Time to opiate use analysis will be performed in Cohort A and Cohort A+B only.

4.2.3.7 Prostate specific antigen (PSA) response

Proportion of patients achieving a PSA response and patients with a confirmed PSA response will be presented with 95% CIs. Best PSA percentage change from baseline will be summarized as continuous variables using descriptive statistics and will be graphically displayed using waterfall plots for Cohort A and Cohort A+B only. In addition percentage change from baseline at Week 12 will be summarized as continuous variables using descriptive statistics and will be graphically displayed using waterfall plots for Cohort A and Cohort A+B only.

For the Week 12 visit, a window of +/- 7 days will be applied to the scheduled visit Day 85. If there is more than one value per patient within the time window, then the closest value to the scheduled visit date will be summarised.

4.2.3.8 Circulating Tumor Cell (CTC) conversion rate

Proportion of patients achieving a CTC conversion will be presented with 95% CIs.

Best change from baseline and best percentage change from baseline in CTC counts will be summarized as continuous variables using descriptive statistics and will be displayed graphically using waterfall plots for Cohort A and Cohort A+B only.

4.2.3.9 Time from randomization to second progression or death (PFS2)

Time from randomization to second progression or death will be analysed using the same methods as in the analysis of the primary endpoint rPFS. The HR and corresponding 95% confidence interval will be based on the Cox model including the stratification variables determined by the pooling strategy as covariates.

A KM plot of time to second progression or death will be presented by treatment group. Summaries of the number and percentage of patients with second progression or death and those who are censored will be provided along with median time to second progression or death for each treatment arm for Cohort A and Cohort A+B only.

4.2.3.10 FACT-P

Summary statistics for FACT-P (FACT-P Total score, TOI, FACT-G, FWB, PWB, PCS and FAPSI 6) score will be presented by treatment group (including means, standard deviation, median and range) for all visits until there are less than the minimum of 20 or 1/3 of patients dosed with evaluable data. Absolute and change from baseline scores for each time point will be calculated for each treatment group.

The proportion of patients with best responses of ‘Improved’, ‘No Change’ and “Worsened” on FACT-P total score and subscales (FWB, PWB, PCS, FAPSI 6, FACT-G) scores including TOI will be compared between treatments using logistic regression with the same methods and covariates as for the logistic regression analysis of pain progression in section 4.2.3.2.

The analysis of FACT-P (FACT-P Total score, TOI, FACT-G, FWB, PWB, PCS and FAPSI 6) will be performed in Cohort A and Cohort A+B only.

FACT-P compliance (overall compliance and by visit compliance) will be summarised for each treatment group

Supportive analyses

Supportive analyses will be performed for the FACT-P total score, and scales (FACT-G total score, TOI, FAPSI-6, FWB, PWB, PCS). Compliance will be analysed for the FACT-P total only.

Change from baseline in the FACT-P total score, and scales (FACT-G total score, TOI, FAPSI-6, FWB, PCS) will be analyzed using a mixed model for repeated measures (MMRM) analysis of all the post-baseline FACT-P (or equivalent) scores for each visit. The study discontinuation visit and the safety follow-up visit will be excluded from this analysis. Restricted maximum likelihood (REML) estimation will be used. The model will include treatment, visit and treatment by visit interaction as explanatory variables and the baseline FACT-P total score as a covariate, along with the baseline FACT-P total score by visit interaction and the stratification variables prior taxane and measurable disease determined by the pooling strategy described in section 4.2.2. Treatment, visit, treatment by visit interaction, baseline FACT-P total score, baseline FACT-P score by visit interaction, prior taxane and measurable disease will be fixed effects in the model. The treatment by visit interaction will remain in the model regardless of significance.

An unstructured covariance matrix will be used to model the within-subject error and the Kenward-Roger approximation will be used to estimate the degrees of freedom. The following provides sample code for implementing the MMRM analysis:

```
proc mixed data= FACTP method = reml;
class TRT VISIT SUBJECT;
model CHBL = TRT VISIT TRT*VISIT BL BL*VISIT PRIOR MEASURE / s
ddfm=kr;
repeated VISIT / type=UN subject=SUBJECT;
lsmeans TRT / at means pdiff diff alpha=0.05 cl;
```

where TRT is the randomised treatment, VISIT is the visit, CHBL is the change from baseline in the FACT-P total score, BL is the baseline FACT-P total score, PRIOR is prior taxane use collected via IVRS and MEASURE is measurable disease at baseline collected via IVRS.

For the estimation of TRT*VISIT means an additional model will be run using all visits and the following lsmeans statement:

```
lsmeans TRT*VISIT / slice=VISIT pdiff diff alpha=0.05 cl;
```

If the fit of the unstructured covariance structure fails to converge, the following covariance structures will be tried in order until convergence is reached: Toeplitz with heterogeneity, autoregressive with heterogeneity, Toeplitz, and autoregressive.

The adjusted mean estimates and corresponding 95% confidence intervals will be presented by visit for each treatment group.

Finally, exploratory descriptive item analysis based on responses at each visit as appropriate will be summarized for overall impact of treatment side effects (FACT-P item GP5, “I am bothered by my side effects of treatment) in Cohort A and Cohort A+B.

4.2.3.11 Pharmacokinetic analysis

The plasma concentration data will be listed within the clinical study report. The pre-dose and post-dose olaparib plasma concentrations will be summarized by nominal sample time, using summary statistics, as detailed below:

- Number of observations (n)
- n > lower limit of quantification (LLOQ)
- Geometric mean (Gmean, calculated as exp[μ], where μ is the mean of the data on a log scale)
- Geometric coefficient of variation (%GCV, calculated as 100 √ [exp(s^2)-1], where s is the standard deviation of the data on a log scale)
- Arithmetic mean (Amean, calculated using non log-transformed data)
- Standard Deviation (using non log-transformed data)
- %CV (using non log-transformed data)
- Median
- Minimum (min)
- Maximum (max)

Reporting of plasma concentrations that are Below Limit of Quantification (BLQ)

Individual olaparib and abiraterone concentrations below their LLOQ of the bioanalytical assay will be reported as NQ (Not Quantifiable) in the listings with the LLOQ defined in the footnotes of the relevant TFLs. Individual plasma concentrations that are Not Reportable will be reported as NR and those that are missing will be reported as NS (No Sample) in the listings. Plasma concentrations that are NQ, NR or NS will be handled as follows for the provision of descriptive statistics.

Descriptive Statistics

- Any values reported as NR or NS will be excluded from the summary tables and figures.
- At a time point where less than or equal to 50% of the concentration values are NQ, all NQ values will be substituted with the LLOQ concentration, and all descriptive statistics will be calculated accordingly.
- At a time point where more than half (but not all) of the values are NQ, the Gmean, %CV, Amean and standard deviation will be set to Not Calculable (NC). The maximum value will be reported from the individual data, and the minimum and median will be set to NQ.
- If all concentrations are NQ at a time point, the Gmean, Amean, minimum, median and maximum will be reported as NQ, and the standard deviation, and %CV will be reported as NC.
- The number of values above LLOQ (n > LLOQ) will be reported for each time point together with the total number of collected values.

Three observations > LLOQ are required as a minimum for a plasma concentration to be summarised. Two values are presented as a minimum and maximum with the other summary statistics as NC.

4.2.4 Concordance between BICR and investigator assessments

4.2.4.1 Concordance between BICR and investigator assessments for rPFS

Concordance between BICR and investigator assessments for rPFS will be summarised by concordance status (concordant or discordant) and type of concordance (according to timing and type of progression event (RECIST and/or bone scan) by treatment group and overall.

The concordance rate will be derived as the proportion of patients where BICR and investigator have agreed on rPFS status (event or censored).

4.2.4.2 Concordance between BICR and investigator assessments for Overall Radiological Objective Response

Concordance between BICR and investigator assessments for best radiological objective response status (response [CR and PR] or non-response [SD, PD, NE and non-PD] categories) and best radiological objective response type (CR, PR, SD, PD, NE and non-PD), according to RECIST1.1 criteria and bone progression status (according to PCWG3 criteria) will be summarised in a shift table by treatment group and overall for subjects with measurable disease (as determined by BICR).

The concordance rate will be derived as the proportion of patients where BICR and investigator, have agreed on response [CR and PR] or non-response [SD, PD, NE and non-PD].

A listing will be created including concordance between BICR and investigator reviews of overall visit response including progression type separately by BICR and investigator assessments, if responses between BICR and investigator assessments are concordant, summarised by subject, treatment group and visit.

4.2.5 Safety analysis

Safety analyses will be presented using the Safety Analysis Set and will be investigated using descriptive statistics. Safety profiles will be assessed in terms of AEs, vital signs (including BP and pulse rate), laboratory data (clinical chemistry and hematology), and physical examination.

Summaries will be presented for Cohort A+B and scheduled visits only. Any unscheduled assessments will be listed. The baseline value is defined as the latest result obtained prior to the start of IP.

Summaries for patients who switch from investigators choice of NHA to olaparib following objective disease progression will be summarised separately for Cohort A+B.

4.2.5.1 Adverse events

Adverse events will be summarized for Cohort A+B of the study and by treatment group. Separate summaries will be produced for Olaparib post switch from investigator choice.

All AEs, both in terms of MedDRA preferred term and CTCAE grade, will be listed and summarised descriptively by count (n) and percentage (%) and treatment group. MedDRA dictionary will be used for coding. Any AE occurring before olaparib/investigators choice of NHA (i.e., before Study Day 1) will be included in the AE listings, but will not be included in the summary tables (unless otherwise stated). These will be referred to as 'pre-treatment'.

An overall summary of the number and percentage of patients in each category will be presented as will an overall summary of the number of episodes in each category. An overall summary of the number and percentage of patients in each category will be presented for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients.

Frequencies and percentages of patients reporting each preferred term will be presented. Total number of events will also be reported separately.

All reported AEs will be included in listings along with the date of onset, date of resolution (if AE is resolved), investigator's assessment of severity and relationship to study drug. A separate listing will be produced for AEs that are on-going in patients who switch from investigator choice to olaparib at the start date of olaparib.

Summary information (the number and percent of patients by treatment) will be tabulated by system organ class (SOC), preferred term and treatment group for:

- All AEs
- All AEs causally related to olaparib/ investigators choice of NHA
- AEs with CTCAE grade 3 or higher
- AEs with CTCAE grade 3 or higher, causally related to olaparib/ investigators choice of NHA
- AEs with outcome of death
- AEs with outcome of death causally related to olaparib/ investigators choice of NHA
- All SAEs
- All SAEs causally related to olaparib/ investigators choice of NHA
- AEs leading to discontinuation of olaparib/ investigators choice of NHA

- AEs leading to discontinuation of olaparib, causally related to olaparib/investigators choice of NHA
- AEs leading to dose reduction of olaparib/ investigators choice of NHA t
- AEs leading to dose interruption of olaparib/ investigators choice of NHA
- Other significant AEs
- Other significant AEs causally related to olaparib/ investigators choice of NHA

Key patient information tables will be produced for:

- AEs causally related to olaparib
- AEs with outcome of death
- All SAEs
- AEs leading to discontinuation of olaparib/ investigators choice of NHA
- Other significant AEs

In addition, a truncated AE table of most common AEs, showing all events that occur in at least 5% of patients overall will be summarised by preferred term, by decreasing frequency for Part B only. This cut-off may be modified after review of the data.

Each AE event rate (per 1000 patient years) will also be summarised by preferred term within each system organ class. For each preferred term, the event rate will be presented and will be defined as the number of patients with that AE divided by the sum of the duration of therapy (for patients without the event) and the time to the AE (for patients with the event) in each group multiplied by 1000.

Adverse events will be assigned CTCAE grades (National cancer institute CTCAE version 4.03) and summaries of the number and percentage of patients will be provided by maximum reported CTCAE grade, SOC, preferred term and actual treatment group. Fluctuations observed in CTCAE grades during study will be listed.

AEs which started prior to first dose or > 30 days following date of last dose will be listed only.

Deaths

A summary of deaths will be provided with number and percentage of patients by cohort and actual treatment group categorised as:

- Related to disease under investigation,

- AE outcome=death,
- Both related to disease under investigation and with AE outcome=death,
- Unrelated to AE or disease under investigation
- Deaths > 30 days after last treatment dose, related to disease under investigation,
- AE with outcome=death > 30 days after last treatment dose,
- Deaths > 30 days after last treatment dose, related to AE or disease under investigation
- Deaths > 30 days after last treatment dose, unrelated to AE or disease under investigation
- Patients with unknown reason for death, and
- Other deaths (not captured above)

Causally related adverse events with an outcome of death will be summarised for the number and percentage of patients by SOC, preferred term, cohort and actual treatment group.

Causally related serious adverse events will be summarised for the number and percentage of patients by SOC, preferred term, cohort and actual treatment group.

Adverse events leading to discontinuation of olaparib/ investigators choice of NHA will be summarised for the number and percentage of patients by SOC, preferred term, cohort and actual treatment group for.

Causally related adverse events leading to discontinuation of olaparib/investigators choice of NHA will be summarised for the number and percentage of patients by SOC, preferred term, cohort and actual treatment group.

In addition, AEs with outcome of death, SAEs, OAEs, AEs leading to discontinuation of treatment and AEs causally related to olaparib/investigators choice of NHA will be listed in key patient information tables.

Listings of AE data will also be produced.

Adverse events of special interest (AESI)

Preferred terms used to identify adverse events of special interest (as defined in section 3.5.3) will be listed before DBL and documented in the Study Master File. Grouped summary tables of certain MedDRA preferred terms will be produced and may also show the individual preferred terms which constitute each AESI grouping. Groupings will be based on preferred

terms provided by the medical team prior to DBL, and a listing of the preferred terms in each grouping will be provided.

Summaries of the above-mentioned grouped AE categories will include number (%) of patients who have:

- At least one AESI presented by outcome
- At least one AESI causally related to study medication
- At least one AESI leading to discontinuation of study medication

A summary of total duration (days) of AESI will be provided for events which have an end date and will be supported by summaries of ongoing AESIs at death and, separately, at data cut-off.

Summary of long term tolerability

To assess long term tolerability, provided that there are a sufficient number of patients with events to warrant it, prevalence plots, life table plots and cumulative incidence plots will be presented for each of the AESI grouped terms and any other events considered important after review of the safety data, provided there are ≥ 10 events.

A prevalence plot provides information on the extent to which the events may be an ongoing burden to patients. The prevalence at time (t) after first dose of study treatment is calculated as the number of patients experiencing the event divided by the number of patients receiving study treatment or in safety follow-up at time t; generally, t is categorised by each day after dosing. The prevalence is plotted over time split by treatment arm. Multiple occurrences of the same event are considered for each patient but a patient is only counted in the numerator whilst they are experiencing one of the occurrences of the event. These plots will only be produced for AESIs that have ≥10 events.

A life table plot can be used to describe the time to onset of the event and specifically when patients are at most risk of first experiencing the event. The hazard, or in other words, the probability of having an AE in a specified time period (e.g. 0-1 months, 1-3 months, 3-6 months, etc.) given that the patient reaches that time period without having an event is plotted for each time period split by treatment. These plots will only be produced for AESIs that have ≥10 events.

A cumulative incidence plot is a plot of the raw cumulative incidence over time. The raw cumulative incidence is the actual probability that a patient will have experienced their first occurrence of the event by a given time point. These plots will only be produced for AESIs that have ≥10 events.

4.2.5.2 Laboratory assessments

Laboratory data (clinical chemistry and haematology) will be summarized. Shift tables will be provided for select tests, where shift from baseline to the worst value within the study will be summarized. Laboratory data outside the reference ranges will be indicated.

For all continuous laboratory assessments, absolute value, change from baseline and percentage change from baseline will be summarised using descriptive statistics at each scheduled assessment time by actual treatment group. For categorical laboratory assessments, shift from baseline will be summarised using frequency and proportion at each scheduled assessment time by actual treatment group.

Shift tables for laboratory values by worst CTCAE grade will be produced, within each part of the study and overall, and for specific parameters separate shift tables indicating hyper- and hypo- directionality of change will be produced. For parameters with no CTCAE grading, shift tables from baseline to worst value on-treatment will be provided (i.e., on-treatment is defined as data collected up until the last dose of olaparib/ investigators choice of NHA). A scatter plot of alanine aminotransferase (ALT) versus total bilirubin, both expressed as multiples of upper limit of normal range, will be produced. The scatter plot will be repeated for aspartate aminotransferase (AST) versus total bilirubin.

Shift tables and plots will be repeated for olaparib post switch from investigator choice, where the baseline is defined as the last observation before the first dose of olaparib (generally the measurement at the study treatment discontinuation visit).

4.2.5.3 Vital signs

Vital signs, including BP (mmHg), body temperature (°C), pulse (beats/minute) and weight (kg), will be summarized at baseline for Cohort A+B by treatment group, including a separate summary for patients who switch from investigators choice of NHA to olaparib. The baseline value is the last pre-dose assessment.

4.2.5.4 Exposure

Summaries of duration of exposure and cumulative exposure over time will be produced for Cohort A+B by treatment group.

The number of patients with study drug reductions, interruptions, or discontinuation and the reasons, will be summarised for Cohort A+B by treatment group.

These summaries will be repeated for the subset of patients who switch from investigators choice of NHA to olaparib.

These data will also be listed.

Summary statistics (mean, standard deviation, median, quartiles, minimum, maximum) will be presented for RDI and PID.

4.2.6 Concomitant medications

Concomitant medications will be summarized by the coded terms. The number of patients receiving a medication will be summarized for Cohort A+B FAS by treatment group. A medication taken from the start of the screening part 2 and onwards is considered concomitant. A patient is only counted once if receiving the medication more than once.

Disallowed medications will be summarised and listed.

4.2.7 Supportive analysis

4.2.7.1 Resource Use

An exploratory health economic analysis of the frequency of metastatic prostate cancer related palliative interventions, time to interventions, and reason for the intervention will be undertaken. In addition, length of stay, ICU use, concomitant medications and analgesic use will be examined.

These analyses will examine the impact of disease and treatment on resource use to primarily support the economic evaluation of olaparib in castrate resistant metastatic prostate cancer.

Resource use may be reported outside of the CSR.

4.2.7.2 Patient reported outcomes (PRO)

The Patient reported outcome endpoints, BPI-SF (Pain Severity domain), BPI-SF (Pain interference domain), Pain Palliation, AQA score, FACT-P (FACT-P Total score, FACT-G total score, TOI, FWB, PWB, PCS and FAPSI 6 and exploratory endpoints [EWB, SWB]), PRO-CTCAE and PGIC are described in Section 3.

The PRO endpoints that are continuous will be summarised using means, standard deviations, medians and ranges by treatment group and at each visit until there are less than the minimum of 20 or 1/3 of patients dosed with evaluable data. For the FACT-P total score, FACT-G total score, TOI, FAPSI-6, PCS, FWB, PWB, EWB, SWB, BPI-SF item #3 (worst pain in 24 hours), pain severity (BPI-SF pain severity domain), pain interference and AQA score, absolute and change from baseline scores for each time point will be calculated for each treatment group. The absolute scores and change from baseline scores will also be summarised descriptively based on plots of unadjusted mean.

Responses on the PRO-CTCAE and PGIC will be summarised descriptively as number of patient and corresponding percentages for each category in the questionnaire at each visit by treatment group. PRO-CTCAE will be summarized for only patients in countries (Argentina Australia, Austria, Canada [English-speaking sites], Germany, Japan, Spain, United Kingdom and United States) where the questionnaire was administered.

Time to pain severity progression and the proportion of patients with pain severity will be analysed using the same methodology as in time to pain progression without any adjustments for multiplicity. This will be repeated for patients who are non-opiate users at baseline.

Proportion of patients with pain palliation will be summarized with corresponding 95% confidence intervals. This will also be compared between olaparib and investigator choice using the same logistic regression model as pain progression (section 4.2.3.2). This will be repeated patients who are non-opiate users at baseline.

Time to deterioration in FACT-P (FACT-P Total score, FACT-G total score, TOI, FWB, PWB, PCS and FAPSI 6) will be analysed using the same methodology as in time to pain progression without any adjustments for multiplicity. Proportion of patients with deterioration in FACT-P scores will be summarized with corresponding 95% confidence intervals. The HRs and 95% CIs will be presented on a forest plot.

Summary measures of overall compliance and compliance over time will be derived separately for BPI-SF, FACT-P, PGIC, PRO-CTCAE.

Supportive analysis for BPI-SF (average BPI-SF worst pain [Item 3], average pain severity score and average pain interference score) will be analyses using the same methodology as the supportive MMRM described in section 4.2.3.10 for FACT-P.

The analysis of PRO endpoints will be performed in Cohort A and Cohort A+B only. The analysis will not include data after patients have received subsequent anti-cancer therapy.

4.2.7.3 EQ-5D-5L

Descriptive statistics will be reported for health state utility index values and visual analogue scale by visits as well as change in these scores from baseline. To support future economic evaluations of olaparib, additional appropriate analyses may be undertaken, for example, mean health state utility pre- and post-treatment, and pre- and post- progression. Further details will be outlined in the payer analysis plan.

EQ-5D-5L may be reported outside of the CSR.

4.2.8 Demographic and baseline data

The following collected data will be listed.

- Patient disposition (received treatment and completed the study)
- Important deviations
- Inclusion in analysis populations
- Demographics (age, age group, race and ethnicity)

The following will be summarized for the full analysis set for Cohort A, B and A+B:

- Patient disposition
- Important protocol deviations
- HRR gene mutations (single gene mutations and co-mutations will be summarized)
- Stratification factors according to the IVRS/IWRS and eCRF
- Inclusion in analysis populations

- Demographics (age, age group, race and ethnicity)
- Patient characteristics at baseline (weight)
- Patient recruitment by country and centre
- Previous disease-related treatment modalities
- Previous disease-related chemotherapy treatments
- Previous/current/post treatment radiotherapy
- Disease characteristics at baseline
- Extent of disease at baseline
- Demographic characteristics
- Time from most recent disease progression to randomisation
- Post-discontinuation disease-related anticancer therapy
- Past/current medical history
- Past medical history of opioid use
- Relevant surgical history at baseline

The following will be summarized for confirmed FMI F1CDx patients and confirmed myriad *gBRCA*m patients for Cohort A, B and A+B:

- Patient disposition
- HRR gene mutations (single gene mutations and co-mutations will be summarized)
- Inclusion in analysis populations
- Demographics (age, age group, race, ethnicity and weight)
- Disease characteristics at baseline

The following will be summarized for patients who switch from investigators choice of NHA to olaparib for Cohort A+B:

- Patient disposition

5. INTERIM ANALYSES

An interim analysis of OS will be performed at the time of the primary rPFS analysis in Cohort A (approximately 35 months after first patient randomized into the study). Approximately 117 deaths (49% maturity) are expected to be accrued in Cohort A at the time of interim OS analysis. The alpha spending for the OS analysis is described in 4.2.3.3.

This study will use an external Independent Data Monitoring Committee (IDMC) to perform interim reviews of accumulating study safety data. The IDMC will not be involved in the interim analysis of OS because it is at the time of the primary rPFS analysis when the study will be unblinded. This committee will be composed of two therapeutic area experts and a statistician, who are not employed by AZ, and do not have any major conflicts of interest. Following the review, the IDMC will recommend whether the study should continue unchanged, be terminated, or be modified in any way. Once the IDMC has reached a recommendation, a report will be provided to AstraZeneca. The report will only include the

recommendation and any potential protocol amendments and it will not contain any unblinded information or reference to the confidential considerations of the committee to have led to their recommendation. A separate IDMC charter will be developed which will contain details of the IDMC members and clearly define the responsibilities of the IDMC.

In addition to the periodic review of safety data by an IDMC, the safety of all AstraZeneca clinical studies is closely monitored on an on-going basis by AstraZeneca representatives in consultation with the Patient Safety Department. Issues identified will be addressed; this could involve, for instance, amendments to the study protocol and letters to investigators.

6. CHANGES OF ANALYSIS FROM PROTOCOL

Section 5.3.1.2 of the protocol details the FACT-P subscales and derivations. The SAP includes FACT-G in addition to the subscales outlined in the protocol.

Section 8.3 of the protocol defines the safety analysis set. An additional safety analysis set has been added to the SAP for patients who switch from investigators choice of NHA to olaparib.

Section 8.4.1.1 states that subjects who have not progressed (defined as CR, PR or SD by RECIST 1.1 for soft tissue disease, or non-PD for bone disease, see Section 5.1.2 and Appendix E) at the time of analysis will be censored at the time of the latest date of their last evaluable RECIST assessment or bone scan. However, if the subject progresses or dies after 2 or more missed radiologic assessments, the subject will be censored at the time of the latest evaluable RECIST 1.1 or bone scan assessment prior to the two missed visits. The censoring approach has been updated in the SAP to take the earliest date of their last evaluable RECIST 1.1 assessment (taking the latest target lesion, non-target lesion or new lesion scan date) or bone scan assessment that showed Non-PD.

Section 8.4.2.2 of the protocol defines 2 consecutive visits as being separated by 3-4 weeks. This has been updated in the SAP to state 2 consecutive visits (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit).

Section 8.4.2.4 of the protocol details that patients who have not experienced any symptomatic skeletal–related event will be censored at time of death, or time of analysis if the patient is living. This has been updated in the SAP to censor the patients at time of death, or time of last SSRE assessment.

Section 8.4.2.5 of the protocol details duration of response will be defined as the time from the date of first documented response (by BICR using RECIST 1.1 and PCWG3) until date of documented progression (by BICR) or death in the absence of disease progression. The text has been updated in the SAP to show patients must have a confirmed response.

Section 8.4.2.10 of the protocol defines 2 consecutive visits as being separated by 3-4 weeks. This has been updated in the SAP to state 2 consecutive visits (with at least 2 weeks between the end of the initial visit and the start of the subsequent visit).

Section 8.4.2.12 of the protocol displays the definition of visit response for FACT-P, FAPSI-8, TOI, PCS and FWB. The SAP has updated the criteria in this table and added FACT-G.

Section 8.4.2.13 of the protocol specifies the confirmation visit for pain palliation must be at least 3 weeks later. This has been updated in the SAP to 2 weeks instead of 3 weeks.

Section 8.2.5 of the protocol specifics demographic and baseline characteristics will be summarized using the full analysis set, for Cohorts A and B of the study, respectively. This has been updated in the SAP to include Cohort A+B.

Section 8.5.8 of the protocol details adverse events and laboratory data will be summarised for Cohort A, Cohort B and Cohort A+B. This has been updated in the SAP to be produced for Cohort A+B only.

Section 8.5.8.4 of the protocol details concomitant medications will be summarised for Cohort A, Cohort B and Cohort A+B. This has been updated in the SAP to be produced for Cohort A+B only.

Section 8.5.8.7 of the protocol details listings and summaries for compliance and exposure will be produced for Cohort A, Cohort B and Cohort A+B. This has been updated in the SAP to be produced for Cohort A+B only.

7. REFERENCES

Burman et al 2009
Burman CF, Sonesson C, Guilbaud O. A recycling framework for the construction of Bonferroni based multiple tests. Stat Med 2009;28:739-61.

Cella et al 1993
Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol. 1993 Mar;11(3):570-9.

Cella et al 1994
Cella, D.F. Manual: Functional Assessment of Cancer Therapy (FACT) Scales and Functional Assessment of HIV Infection (FAHI) Scale, Chicago: Rush-Presbyterian-St Luke's Medical Center. 1994.

Cella et al 2009
Cella D, Nichol MB, Eton D, Nelson JB, Mulani P. Estimating clinically meaningful changes for the Functional Assessment of Cancer Therapy-Prostate: results from a clinical trial of patients with metastatic hormone-refractory prostate cancer.
Value Health 2009; Jan-Feb;12(1):124-9.

Chung et al 2014
Chung KC, Barlev A, Braun AH, Qian Y, Zagari M. Assessing analgesic use in patients with advanced cancer: development of a new scale--the Analgesic Quantification Algorithm. Pain Med 2014;15:225- 32.

Esper et al 1997
Esper P, Mo F, Chodak G, Sinner M, Cella DF, Pienta KJ. Measuring quality of life in men with prostate cancer using the Functional Assessment of Cancer Therapy-Prostate instrument. Urology 1997;50:920-8.

Pintilie
Pintilie M. Competing risks: A practical perspective. Wiley.

Robins et al 1991
Robins JM, Tsiatis AA. Correcting for noncompliance in randomised trials using rank preserving structural failure time models. Commun Stat Theory Methods. 1991; 20:2609-2631.

Robins 1993
Robins J. Information recovery and bias adjustment in proportional hazards regression analysis of randomised trials using surrogate markers. Proceedings of the Biopharmaceutical Section, ASA. 1993; 24-33.

Sun and Chen 2010
Sun X, Chen C. Comparison of Finkelstein's Method With the Conventional Approach for IntervalCensored Data Analysis. Stat Biopharm Res. 2010; 2:97-108

8. APPENDIX (NOT APPLICABLE)